# New approaches to improving the accuracy and outcome of Intensity Modulated Radiotherapy

## Thesis

# New approaches to improving the accuracy and outcome of Intensity Modulated Radiotherapy

by

Ian Richard Cowley M.Sci. M.A. (Cantab)

A thesis submitted to the

Faculty of Life Sciences of the

Open University

for the degree of Doctor of Philosophy.

Supported by a sponsoring establishment:

Addenbrooke's Hospital, Cambridge

Date of submission: 1st February 2006

PAGES NOT SCANNED AT THE REQUEST OF THE UNIVERSITY

SEE ORIGINAL COPY OF THE THESIS FOR THIS MATERIAL

# Abstract

Several aspects of the IMRT treatment chain are evaluated with a view to improving the accuracy with which the treatment can be delivered and also the radiobiological outcome for the patient. A safe, non-invasive way of evaluating patient repositioning for multi-fraction treatments is developed and presented, with results highlighting the head-and-neck immobilisation system that provides the best results at Addenbrooke's radiotherapy centre. Previously published dose evaluation metrics are studied in depth, identifying ways in which they differ. A new method of calculating the popular Gamma index is proposed, along with a new Kappa index which can highlight areas of dose mismatch more readily than the Gamma index. Using a specially developed dose calculation system, systematic errors found with the IMRT treatment machines at Addenbrooke's are simulated to determine the dose differences created by these errors, which are then evaluated using the Kappa index with typical tolerances used by previous authors. The concept of equivalent uniform dose (EUD) is then used to assess whether the tolerances for distance and dose are equivalent to each other. They are found not to be equivalent, and better relative values of the two quantities are suggested for use in composite indices. The use of the EUD is extended further into analysing the equivalent dose differences of the treatment delivery errors previously calculated. It is found that the equivalent dose to the target volume is changed by more than the composite indices indicate. The EUD is then used to calculate normal tissue complication probabilities, the probability of uncomplicated control and therapeutic gain. It is found that the dose errors predicted can actually be beneficial to the patient outcome, although not always. It is concluded that dose evaluation for complex treatments such as IMRT should be more radiobiologically-based in order to assess patient outcomes properly.

# Contents

## Acknowledgements

Firstly, my thanks go to Siemens, who originally provided the money to Addenbrooke's to fund two-thirds of this research without ever forcing me into any particular topic. Addenbrooke's Hospital, the sponsoring establishment, provided the rest of the funds, and agreed to fund me beyond my nominal finish date.

Thanks clearly go to my supervisors, Simon Thomas, Philip Dendy and Sarah Brooke, especially the latter two who supervised me in their retirement and pregnancies respectively. Your encouragement and pointing out of my split infinitives has enabled me to happily hand in a polished product.

My thanks go to the radiotherapy physics department at Addenbrooke's, the members of which are too numerous to mention explicitly. Special thanks to Nikki Twyman for teaching me the joys of Nikki Physics, Hannah Chantler for helping with the laser scanning in chapter 2 and Lindsay Hodgson, my officemate who kept me sane through my three-plus years. Couldn't have done it without you, Linds.

And finally to Anna.
You are my world. This PhD is for you.

# List of abbreviations

3D-CRT: Three-dimensional conformal radiotherapy

C(A)T: Computed (axial) tomography

CRC: Cyclic redundancy check

CTV: Clinical target volume

DICOM: Digital imaging and communications in medicine

DICOM-RT: Radiotherapy extensions to the DICOM standard

dpi: Dots per inch

DRR: Digitally reconstructed radiograph

DTA: Distance to agreement

EPID: Electronic portal imaging device

EUD: Equivalent uniform dose

GTV: Gross tumour volume

ICE: Independent calculation engine

ICRU: International commission on radiation units and measurements

IMRT: Intensity modulated radiotherapy

Linac: Linear accelerator

MLC: Multi-leaf collimator

MSF: Multiple static field IMRT

MU: Monitor unit

NTCP: Normal tissue complication probability

OAR: Organ at risk

PETG: Glycolised polyethylene terephthalate

PI: Portal image

PMMA: Polymethyl methacrylate (Perspex™)

PTV: Planning target volume

PUC: Probability of uncomplicated control

PVC: Polyvinyl chloride

QA: Quality assurance

SSD: Source-surface distance

TCP: Tumour control probability

$TD_{50}$: The dose required to give a 50% TCP or NTCP

TG: Therapeutic gain

V&R: Verify and record

# Table of illustrations

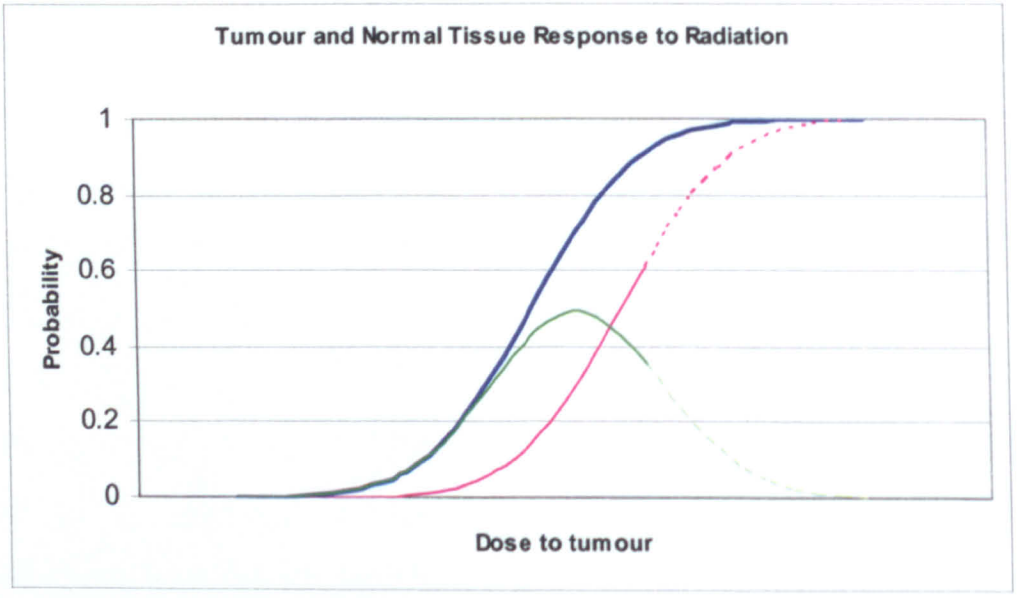



**Figure 1.1**: Schematic representation of tumour control and normal tissue complication. The blue line shows tumour control, the pink line shows normal tissue complication, and the green line shows the probability of uncomplicated control. The dotted lines indicate extrapolation of data, as irradiation of healthy tissue to this level has not been researched fully.

# Chapter 1 – Introduction

Radiotherapy is the use of ionising radiation for therapeutic purposes, aiming to irradiate a target cancer tumour volume while minimising the dose given to surrounding healthy tissue and the related side-effects. Radiotherapy takes two forms, palliative and radical. The former provides pain relief and improved quality of life for patients with advanced disease, while the latter's primary goal is curative control of the tumour. Radiotherapy can be given using external beams of ionising radiation (the most common use of the word radiotherapy), implanted or inserted radioactive materials (brachytherapy) or unsealed radioactive sources inside the body. External beam radiotherapy typically uses linear accelerators (linacs) to produce high-energy (between 4 and 25 MV) X-ray and electron beams and X-ray tubes to produce low-energy (<500kV) X-ray beams. Gamma rays produced from radioactive materials may be used, such as those from cobalt-60, which emits 2 mono-energetic $\gamma$-rays of energies close to 1.2 MeV. The work presented here focuses solely on external-beam radiotherapy using linear accelerators in photon mode at energies of 6 MV and 15 MV.

## Origins of the use of radiation for therapy

Radiotherapy began in 1896 within a year of the discovery of the X-ray (Freund 1897). Doctors understood from an early point in the history of radiotherapy the need to localise radiation at the target as much as possible due to the side-effects of radiation on healthy tissue. Holthusen (1936) first described the shape of the curve showing the response of both tumour and healthy tissue to radiation and figure 1.1 shows the sigmoidal shape he described. Holthusen also identified the healing probability, that is the probability of uncomplicated control, given by $H=T\times(1-N)$ where T is the tumour control probability and N is the normal tissue complication probability. Other similar

predictors of the outcome of radiotherapy are the Therapeutic Ratio, given by TR=T/N and the Therapeutic Gain, given by TG=T-N. The probability of tumour control without complication is typically quite small, so oncologists may choose to maximise the Therapeutic Ratio when prescribing treatments and setting limits on normal tissue doses for radiotherapy. The response curves of normal tissue and cancerous tissue are very similar in shape and the probability of uncomplicated control if both are irradiated to the same dose is effectively zero. However, the dose axis on the graph in figure 1.1 represents the dose delivered to the tumour. By localising dose to the target and irradiating surrounding tissue to a lower dose, the normal tissue curve is moved to the right, away from the tumour control curve. The reduction of dose to normal tissue contributes to the therapeutic ratio as well as other factors not directly related to dose.

Other factors that may move the radiation response curves relative to each other include differences in oxygenation of the tumour, the aggressiveness of the cancer, the location of the tumour, the radiosensitivity of nearby normal tissue and fractionation. Fractionation involves dividing radiotherapy treatment of the patient into several sessions, given daily over several weeks. Since some healthy tissues can recover from radiation faster than diseased tissue, fractionation helps to further separate the normal tissue and tumour curves on the response graph. However, the main therapeutic advantage comes from reducing as far as possible the dose to normal tissue while maintaining a high target dose.

**Development of radiotherapy**

Prescription of radiation dose given to a specific point in a patient can be calculated very easily for one or two incident X-ray beams, but using a single beam or two parallel opposed beams means that all tissue in the path of the radiation before and beyond the tumour may receive as much dose as the target tumour. Until the mid-1970s

radiotherapy treatment plans were typically based on orthogonal pairs of radiography images, and were treated using beam cross-sections that were geometrically simple shapes such as circles or rectangles, the dosimetry of more complex shapes being too difficult to calculate by hand or too time-consuming to measure for every patient.

The development of computers through the 1970s allowed automated calculation of dose. Dose distributions were typically calculated as 2-dimensional planes, for a single slice of a patient, due to the speed of computers. Hand-calculated plans were still common, and treatment planners would calculate dose only at a small number of critical points, such as the tumour itself and any particularly radiosensitive organs whose dose needs to be kept as low as possible.

The development of the computed [axial] tomography (CT or CAT) scanner in 1972 (Ambrose and Hounsfield 1973, Hounsfield 1973) allowed treatment planning based on images of axial slices of patient anatomy which showed the tumour location more clearly and allowed oncologists to identify better what to treat and what not to. The rapid development of computers through the 1980s and 1990s and the procurement of more CT scanners led to the ability to calculate full 3-dimensional dose distributions across many slices of CT images. This had a two-fold effect: doses could be tailored much more easily to cover the whole target volume rather than using the extrapolation of a single plane into three dimensions, and thick blocks of metal could be more easily used to conform the beam's cross-section to the target volume shape, thus sparing areas in each beam's-eye-view projection that can be seen not to contain cancer cells.

Current three-dimensional conformal radiotherapy (3D-CRT) thus utilises multiple beam apertures of complex geometry to deliver higher doses to target volumes while sparing more healthy tissue than ever before.

Sparing truly healthy tissue can only be good for the patient, as there can be no benefit to any patient in irradiating tissue that is not involved with the tumour that is

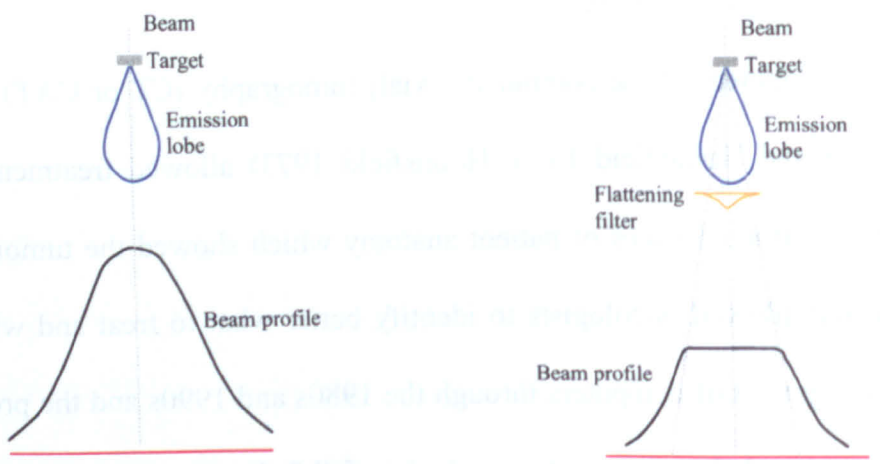



**Figure 1.2:** The X-ray output from a linac target is mainly forwards, the rounded emission lobe creating the beam profile on the left. Use of a specially-shaped flattening filter produces a flat beam profile as shown on the right

being targeted (Suit 2002). Reducing dose to the healthy tissue surrounding a tumour will reduce side-effects and the possibility of secondary cancers induced by the radiation itself. However, the possibility exists that traditional non-conformal radiotherapy may be inadvertently treating and curing microscopic spread of diseased tissue that cannot be imaged or detected. Thus by sparing apparently healthy tissue near the apparent tumour boundary, the treatment may be missing diseased tissue and thereby not completely curing the disease itself. A clinical trial carried out (Dearnaley *et al* 1999) studied the effects on men treated for prostate cancer with traditional or conformal radiotherapy. The results showed that treating conformally gave no significant difference in tumour control but significantly reduced the probability and severity of proctitis and rectal bleeding.

Therefore, the delivery of accurate and beneficial radiotherapy depends critically on accurate definition of the boundaries of the diseased area. One of the problems to be addressed in this thesis focuses on the definition of margins and their use.

**The use of radiation for treatment**

Accurate radiotherapy relies on controlling beam parameters carefully to ensure reproducible and calculable dosimetry. The mainly-forward production of X-rays in a linac produces a beam with a tear-drop shape giving isodose profiles within a patient that are rounded across the beam profile. Linac design therefore incorporates a flattening filter to produce a more useful flat, symmetric beam profile, emulating a top-hat function as closely as possible (see figure 1.2). Flat-profile beams such as these provide much greater flexibility to the radiotherapy treatment planner, who can then use any shape of beam aperture anywhere within the whole cross-section area of the flattened primary beam and know that a similar fluence of radiation will be given everywhere.

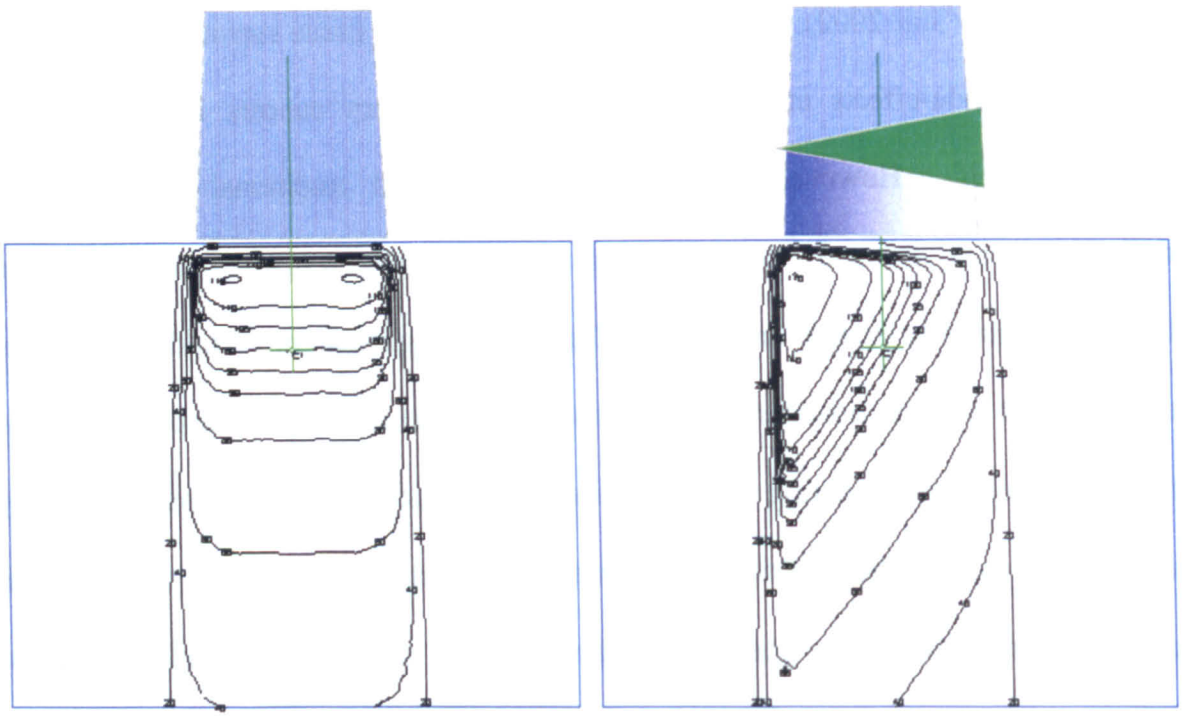


**Figure 1.3:** Isodose (lines of equal dose) profiles created by a flat symmetric beam and a 60° wedged beam. The green triangle represents a physical wedge, with attenuation of the radiation (in blue) shown below it. The sloped isodoses are at 60° to the horizontal, specifically the 50% isodose on the central axis of the beam

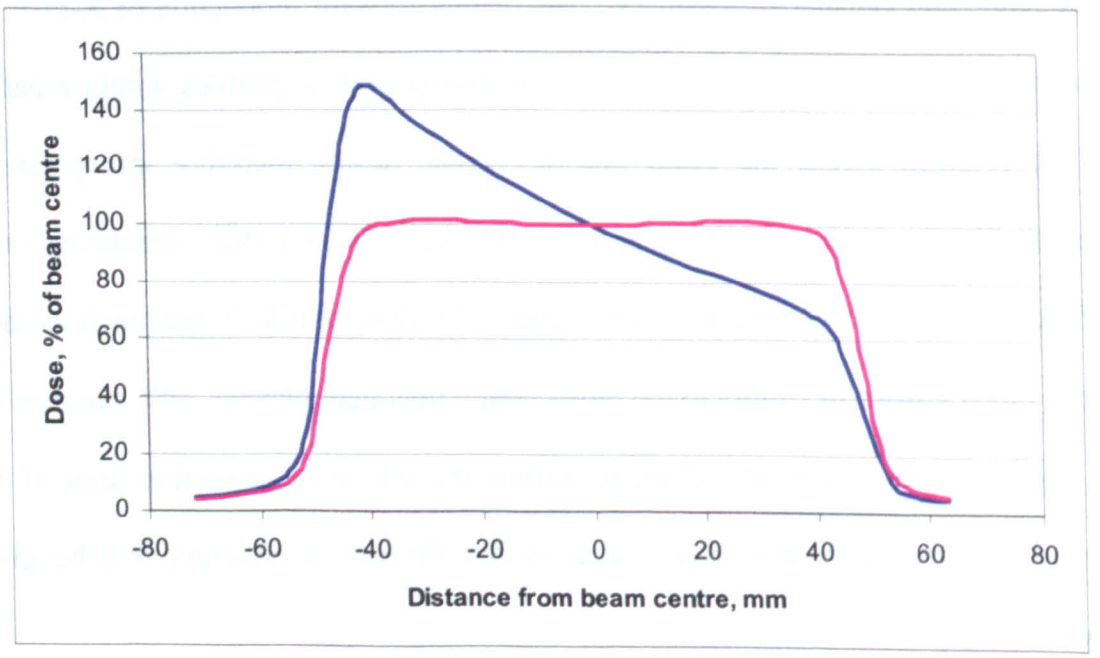



**Figure 1.4:** Flat beam profile and 60°-wedge beam profile for a 10 cm × 10 cm field, measured at 5 cm depth in water. Both profiles are normalised to 100% at their centres.

It is not possible to achieve perfectly flat beam profiles at all locations for all beam aperture shapes, so practically a beam is tuned to give a flat profile for one field size at one depth measured in water, for example a $10\times10$ cm$^2$ field measured at 10 cm depth with the water surface 100 cm from the radiation source. All unmodified beams profiles are designed to be symmetric across their profile when exiting the linac.

Flat symmetric beams, however, provide little scope in radiotherapy for avoiding sensitive structures that may be close to the target volume or for "dose-sculpting" of dose distributions around complex target shapes and intricate sensitive structures. Modulation of the flat beam profile can provide additional tools to the treatment planner. Modification of the intensity of radiotherapy beams comes in many guises, and has been used in its most simple forms for many years. Wedge-shaped beam profiles are the most common form of beam modulation used. A deliberately sloped and asymmetric beam profile is produced to achieve better dose distribution where two beams meet at an angle or to compensate for oblique patient surfaces.

Figure 1.3 shows how wedged profiles can be created by introducing a physical metal wedge into the beam's path to attenuate one side of the beam more strongly than another. More usually now, wedged beams can be created by dynamically closing or opening a beam-collimating jaw during the irradiation (Kijewski *et al* 1978, Levene *et al* 1978, Leavitt *et al* 1990, Beavis *et al* 1996, van Santvoort 1998, Miften *et al* 2000). By allowing more radiation on one side of the beam than the other, the same wedged profile as a physical wedge is created. The radiation response of tissue is sensitive only to total dose given over the short timescale of a treatment beam, so there is direct equivalence between the dynamic and static wedge techniques. Physical metal wedges therefore modulate the intensity of a beam, whereas dynamic wedges modulate the total fluence of constant-intensity radiation given. Wedged beam profiles are described by the angle that the isodose lines they create make with isodose lines for un-wedged

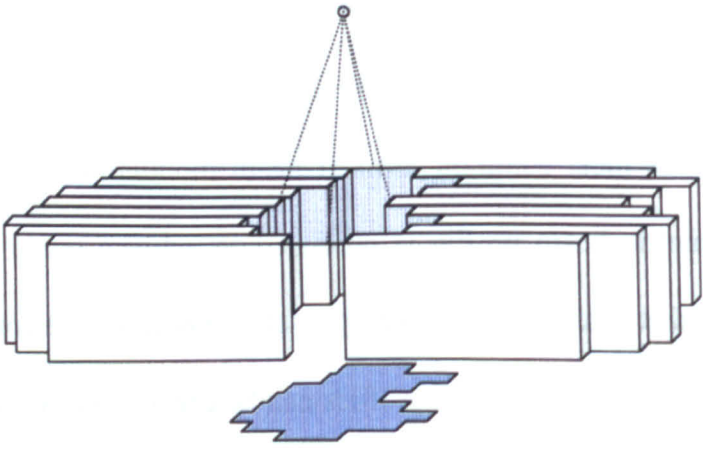


**Figure 1.5:** Opposing banks of MLC leaves can create almost any aperture to shape the radiation beam's cross-section. The leaves are shown as flat-ended for simplicity. The leaves and radiation source distance are not to scale.

profiles, as shown in figure 1.3. Figure 1.4 shows the dose intensity across a 60° wedged beam.

More complicated modulation can be created using metal compensators with varying thickness across the beam cross-section. These compensators work in the same way as a metal wedge by attenuating the beam differentially across the beam aperture. This technique can be used to compensate for unevenness in a patient's outline - by using more intense radiation on thicker parts of a patient and less on thinner parts a more even dose distribution can be achieved at the target depth. Uneven surfaces may alternatively be packed with tissue-equivalent bolus to create a more even surface, underneath which dose distributions will be more uniform.

**Intensity modulated radiotherapy**

The next step in developing beam shaping in radiotherapy is Intensity Modulated Radiotherapy (IMRT), a general term that has come to mean radiotherapy where modulation of the beam fluence is more complex than can be achieved with a collimator, wedge or compensator (Palta and Mackie 2003, Webb 2001). IMRT beam outlines are conformed to the target volume in a similar way to conformal treatments, and thus spare healthy tissue in the same way, but are modulated in the radiation intensity, or more accurately the radiation fluence (Webb and Lomax 2001), across the beam cross-section to vary the way dose is delivered across the target volume. IMRT delivery, in its complexity, is an automated process utilising computer-controlled motorized components of the linac head between radiation source and patient to modulate the beam, rather than the manual inclusion of specially-cut compensators or other blocking devices manually manufactured from low melting point alloy. This complexity is created by the use of multi-leaf collimators (MLCs), opposing banks of thick metal leaves between 1 mm and 0.5 cm wide that can be independently moved

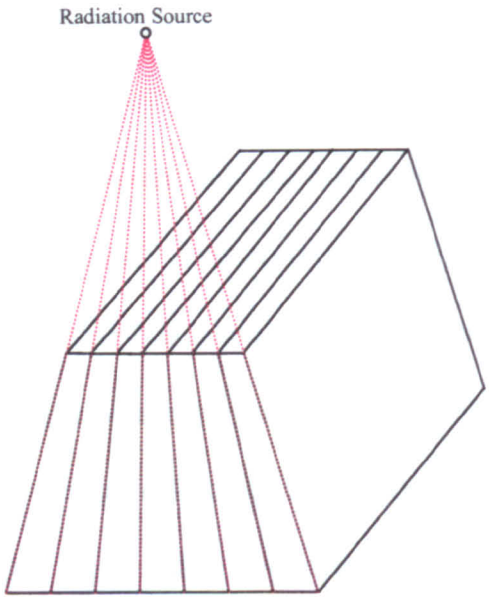



**Figure 1.6:** End-on schematic diagram of MLC leaves, showing the tapered sides that project back to the radiation source. The leaves and radiation source are not to scale.

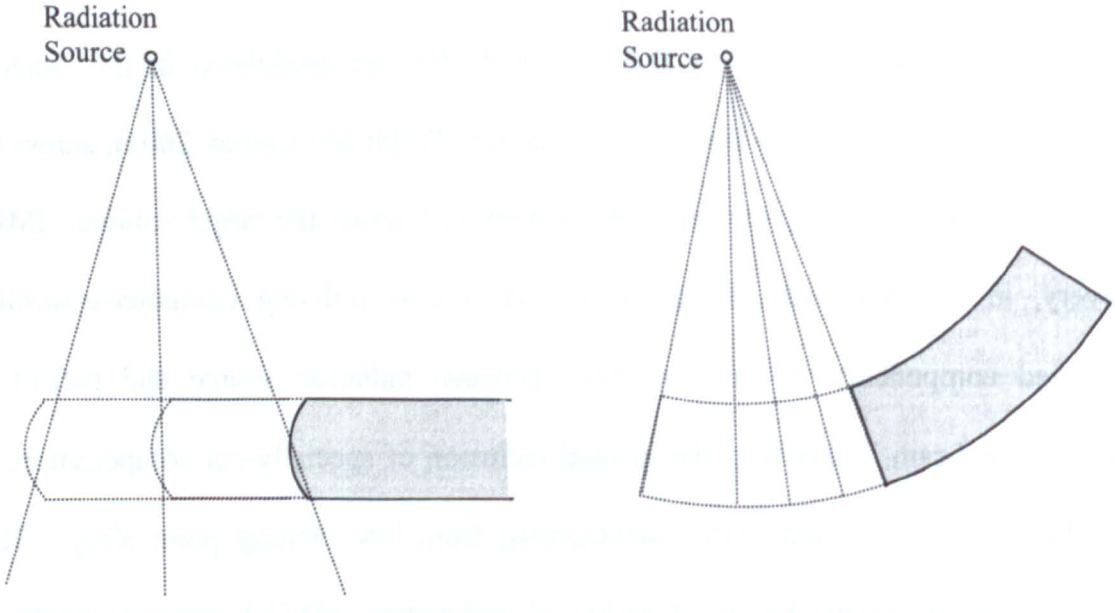



**Figure 1.7:** Schematic diagram of round-ended and flat-ended MLC. Round-ended leaves travel in a plane perpendicular to the beam axis, whereas flat-ended leaves travel in an arc centred on the radiation source. Both methods aim to achieve a constant penumbral width for the radiation beam edge. Diagrams are not to scale.

into or out of the beam to create almost any shape of beam aperture (Carlson 2001), as shown in figure 1.5 (previous page). The MLC leaves are positioned approximately one-third to one-half of the distance from the radiation source to the patient, thus the leaf thickness is projected to between 3 mm and 1 cm width at the patient.

The design of such mobile collimators follows two techniques, single-focused and double-focused. Single-focused leaves have the leaf sides tapered so that the projection of their edges goes back to the radiation source (see figure 1.6). The leaf ends are rounded and the MLC leaves move only in a plane perpendicular (see figure 1.7) to the beam axis, in such a way that the penumbra caused by having a non-uniform thickness of leaf at the radiation field edge is approximately constant across all positions of the leaves. Varian[1] and Elekta[2] produce linacs which include single-focused leaves, as do most add-on devices that can be purchased for fitting to machines sold without MLCs.

Siemens[3] linacs utilise double-focused MLC leaves – these leaves have tapered sides like single-focused leaves, but also have flat ends. The leaves travel in an arc centred on the radiation source such that the flat leaf ends are always parallel to the radiation coming from the source and thus the field edge penumbra is constant for all field sizes (see figure 1.7).

Multi-leaf collimators were initially developed to replace the need for heavy metal blocks to be manufactured and manually inserted to conform beam shapes to target volumes for each beam of every patient. MLC leaves can form almost any shape of beam aperture, albeit with a stepwise edge, and the automatic nature of their positioning helped pave the way for IMRT treatments although their main use is still the replacement of metal blocks.

---

[1] Varian, Palo Alto, California
[2] Elekta, Crawley, UK
[3] Siemens Medical Solutions, Concord, California

Convery and Rosenbloom (1992) first proposed the concept of dynamically-moving MLC leaves to create modulation of a radiotherapy beam. The MLC leaves move across the path of the beam during irradiation and the fluence at any point is determined by the length of time that it is shielded or open to the radiation. Shortly afterwards Nomos[1] announced their MIMiC system for IMRT. This consisted of a radiation slit 2 cm in width whose 1x1 $cm^2$ beam area segments could be open or shielded by collimators. The slit rotates around the patient as the collimators open and close, delivering radiation to a slice of the patient. The couch, with the patient on it, is then moved longitudinally by 2 cm and the process repeats until the target volume has been fully treated. In 1993, Mackie *et al.* proposed a helical version of the Nomos system, in which the patient constantly moves through the machine as the radiation slit rotates. The system is a treatment analogy to the scanning system of CT imaging, and is called tomotherapy. In 1994, Bortfeld *et al.* treated the first patients with sequential static fields of multileaf-collimated radiation to create fluence-modulated beams. This method is called MSF (multiple static field) or step-and-shoot, because the collimators move (step) only when the radiation is off (not shooting). IMRT today is typically delivered using either the dynamic MLC technique or with MSF. Tomotherapy is now becoming more common as complete plan-and-treat systems have become available. The work presented here deals solely with step-and-shoot IMRT using multileaf collimators.

Multileaf collimators introduce many possible errors into the delivery of radiotherapy over the use of static metal collimators and simple beam aperture shapes. The possible effects of these errors, and suggested actions to take against them, must be investigated and an important part of the work presented here deals with these aspects of IMRT.

[1] North American Scientific, Chatsworth, California

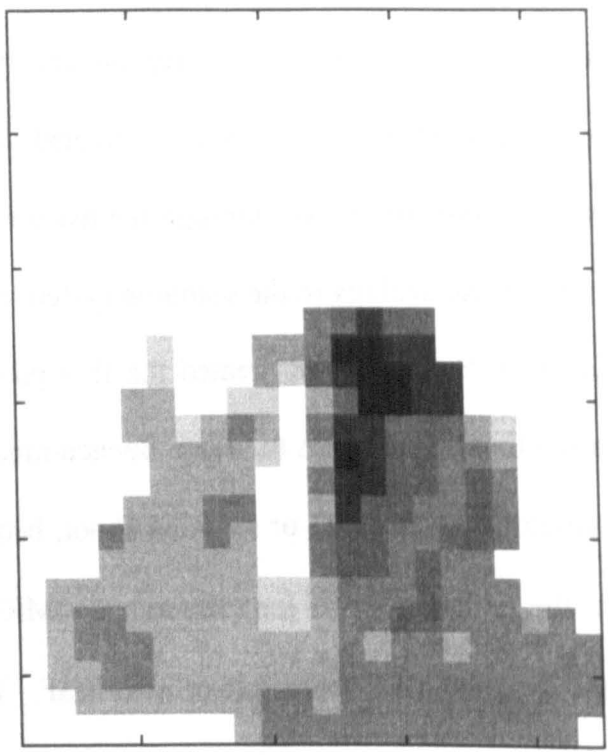


**Figure 1.8:** Example fluence modulation map for an IMRT beam. Black indicates maximum intensity, and in this case there is an area of low intensity in the centre of the beam to help spare dose to the spinal cord.

The complexity of IMRT, with many beams, modulated fluences and complex delivery most often requires the treatment to be automatically calculated by computer in a process called inverse planning. Forward-planned treatments involve the treatment planner deciding on beam angles, aperture shapes and beam doses to construct an acceptable dose distribution within the patient, which is calculated and displayed by a computerised planning system. While this method is suitable for simple geometries of beam aperture with flat or wedged beam profiles, it is effectively a trial and error process of deciding the best beam arrangement, and it is not suitable for highly-modulated complex beams arrangements such as those used for IMRT. In the process of inverse planning the treatment planner usually decides upon suitable beam entry angles, to avoid critical sensitive organs as much as possible, and specifies the desired dose to all organs, both target volume and structures to avoid. Meedt *et al* (2003) described a system for optimizing the beam angles, although this is not yet available in commercial systems. The planning system then attempts to develop the best arrangement of beam apertures and modulation to deliver the desired dose. The process of defining dose and allowing the computer system to work out the arrangement of beam intensities is the inverse of forward planning in which the beams are chosen first and the computer calculates the dose.

Most current IMRT inverse planning systems operate on a basis of optimizing a fluence-modulated beam cross-section. This method of inverse planning starts with a beam outline that is slightly larger than the projection of the target volume, and then finds the best arrangement of modulated fluence within that outline. The resulting modulation pattern, such as that shown in figure 1.8 then needs to be delivered. The process of dividing this pattern into a series of individual subfields or segments is called segmentation. There are several segmentation algorithms that can be used to divide a modulated fluence map into segments of which two of the most common segmentation

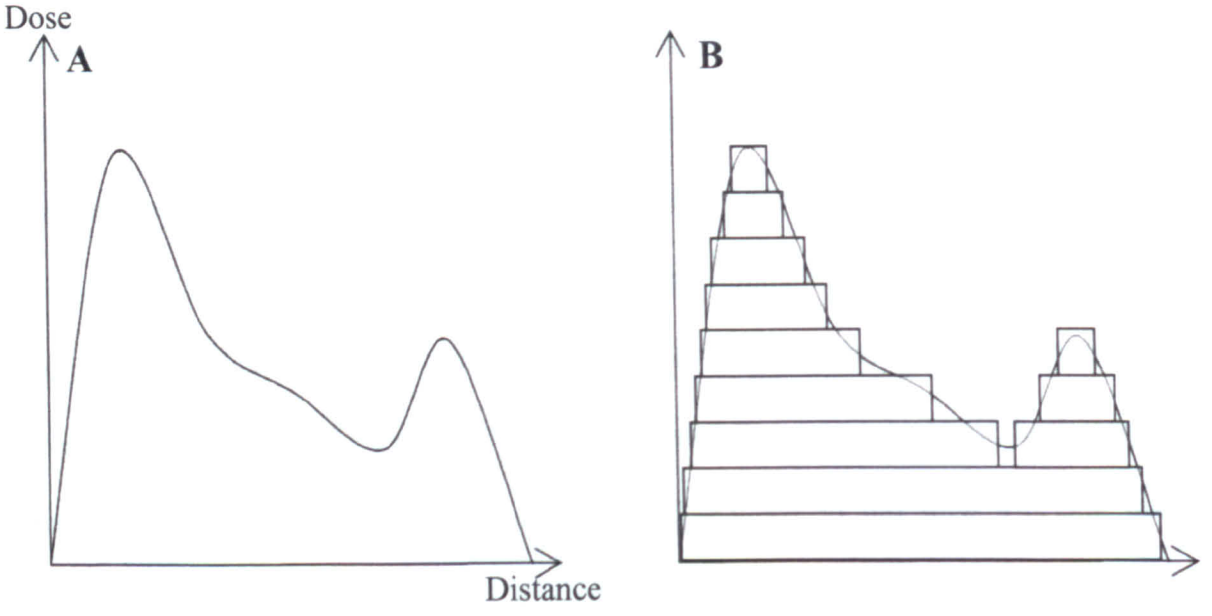



**Figure 1.9:** Schematic example of close-down IMRT segmentation. The 1-D profile (A) the left is split into 9 intensity levels and can be treated with the 12 fields shown

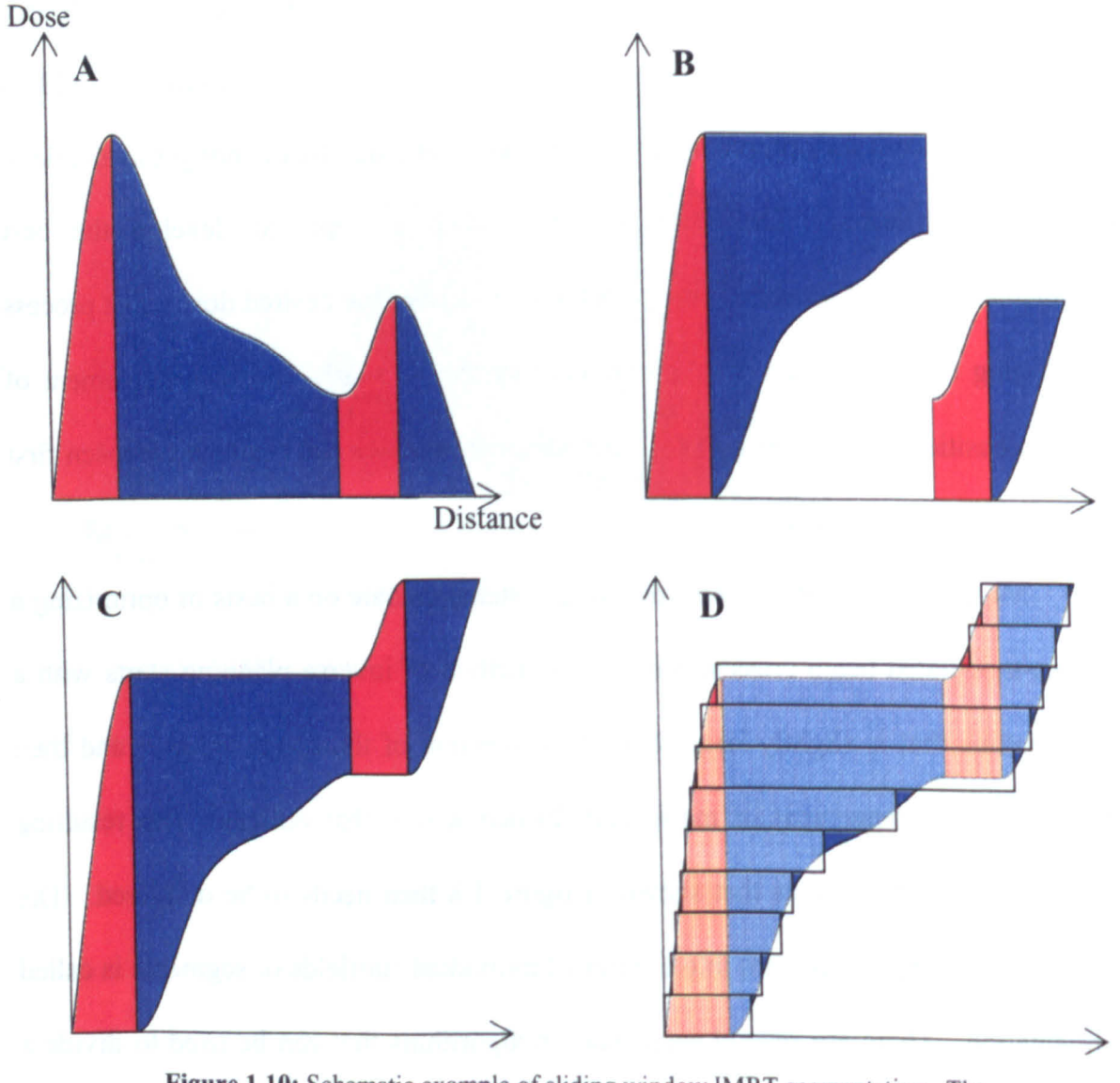



**Figure 1.10:** Schematic example of sliding window IMRT segmentation. The blue sections of the profile (A) are inverted to remove negative gradients (B) , then shifted up to form a continuous profile (C), which can then be divided into segments (D).

algorithms used are the "close-down" method and the "sliding window" technique (Bortfeld *et al* 1994, Siochi 1999). The close-down technique is the easiest to understand and figure 1.9 shows a schematic example of the segmentation of a fluence profile using this method. The curve on the left side (A) shows the desired dose profile in 1 dimension, representing one MLC leaf. The vertical axis shows dose, with the horizontal axis showing position across the irradiated area. The maximum dose scale is divided into equal units and fields fitted to the profile shape, as shown in the right-hand pane B.

Figure 1.10 shows the segmentation of the same dose profile using the sliding window technique. In this method, the MLC leaves are made to travel in only one direction across the field, which speeds up the delivery of the treatment. If there are both negative and positive gradients in the dose profile, then backwards movement of the MLC leaves will be required, so the first step of the segmentation is to remove these gradients. The areas of the profile with negative gradients (shown in blue) are inverted, shown in pane B, and sections of the dose profile are moved up (pane C) to make a continuous positive-gradient profile. This profile is then divided vertically into the same MU units as for the close-down method and the MLC openings are fitted to the profile. Pane D shows that subsequent MLC openings only have the edges of the openings moving from the left to the right, which minimizes MLC motion during beam-off and therefore the treatment time. This method of ensuring that the dose profile has only positive gradients is the method used for plotting the leaf motion of dynamic IMRT as the curve shows the path the MLC leaves need to take to define the original dose profile. Most segmentation algorithms divide the vertical scale into equal units of dose for ease of programming the algorithm. It is clear however that using a variable division would produce a stepped profile that is closer to the desired smooth profile

(Beavis 2001) by using smaller MU segments where the gradient of the desired dose profile is shallow or unchanging.

Figures 1.9 and 1.10 show the segmentation of 1D smooth dose profiles into step-and-shoot MLC openings. The process for segmenting a 2D profile is repeated for each MLC leaf pair and appropriate leaf segments are added together to treat 2D apertures. There are certain constraints on leaf motion and position from different MLC constructions that define appropriate grouping of single MLC openings. Varian and Elekta leaves, for example, have rounded leaf-ends. This means that even if the leaf-ends are touching, radiation is still transmitted through the rounded ends, and the radiation field is thus larger than the positions of the leaf-ends. This means that there is a minimum field size that is achievable. Many systems, including Elekta and Siemens, do not allow leaves to interdigitate. This means that if leaf 2 on one bank of leaves is pulled back from leaves 1 and 3, leaf 2 on the opposite bank is not allowed to follow it and travel between leaves 1 and 3. This is a mechanical constraint designed to prevent the corners of the MLCs colliding, which may prevent them from forming the desired field or damage the driving mechanisms. With Elekta MLCs, this interdigitation constraint also extends forward from the MLC ends due to the minimum approach rule.

Inverse planning systems do not optimize fluence based on continuous smooth profiles, as the computation time would be prohibitively long, and the profile will need be split into segments which will form steps in the profile anyway, as in figure 1.9. Optimization therefore normally creates a grid of discrete fluence pixels, such as that shown in figure 1.8. The sizes of the different pixels are a constraint both of the treatment machine and the planning system. The width of the pixels is defined by the width of the MLC leaves themselves. The length of the pixels is defined by the treatment planning system and/or treatment planner – some planning systems have a

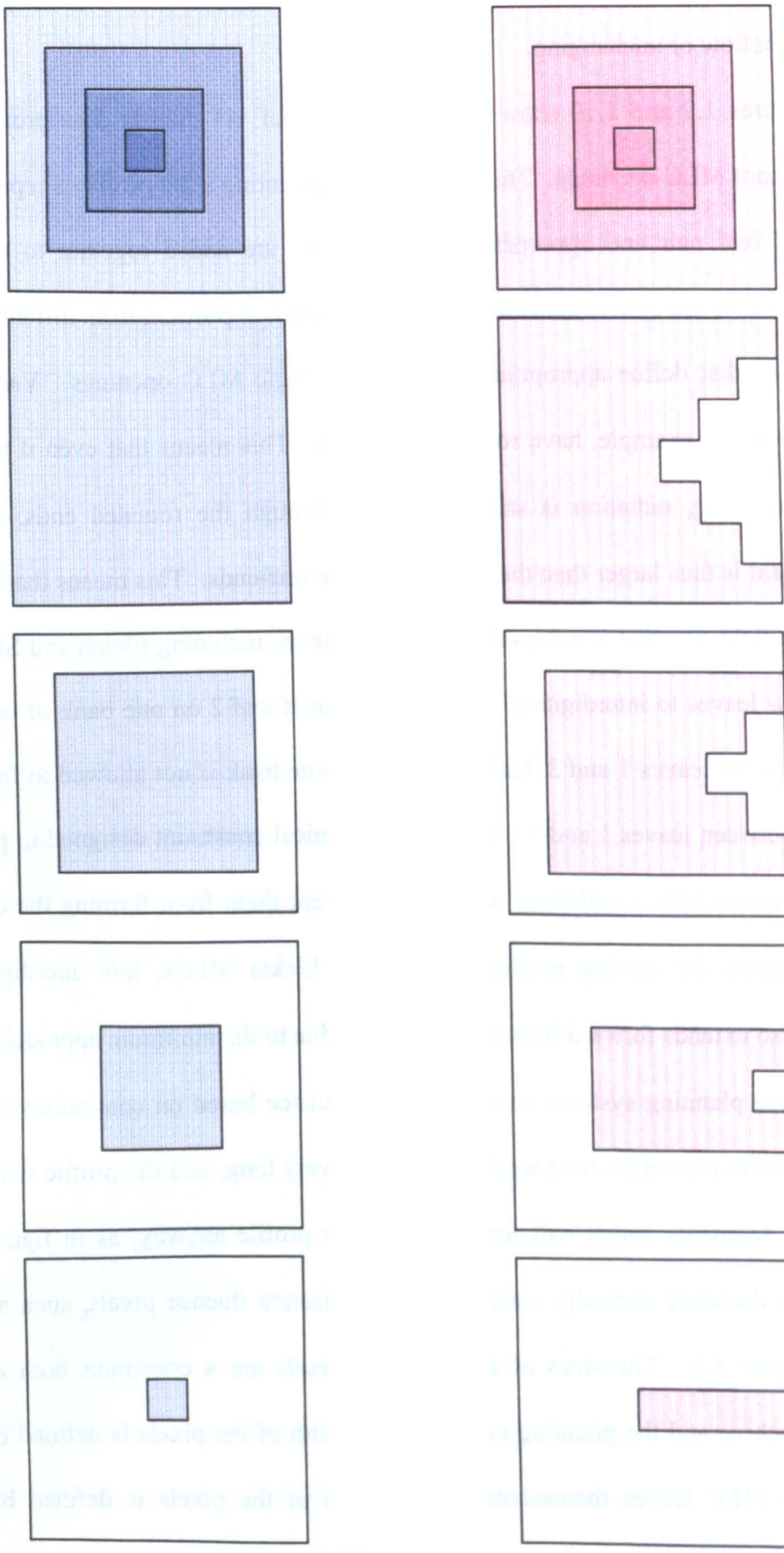


**Figure 1.11:** Different ways of segmenting the same modulated fluence profile. The left-hand column shows segmentation of the pyramid into four square fields in the "close-down" technique. The right-hand column shows segmentation of the same pyramid shape using the "sliding window" technique where all field edges move from left to right. Segments generated by ImFast software.

fixed pixel size. In general, finer sizing of these pixels will result in a stepped profile closer to the desired smooth profile, but will require more segments to treat.

Figure 1.11 shows how a simple modulated pattern defining a pyramid of dose can be segmented using the sliding window and close-down techniques. The left-hand column shows the MLC shapes that can be irradiated to produce the fluence map shown at the top of the column by using the close-down method. Several squares, each smaller than the last, are irradiated to build up the shape. The sliding window technique, shown in the right-hand column, does not use square fields but is able to build up the same overall fluence of radiation.

IMRT can deliver much more complex dose distributions than can 3D-CRT and thus allows the delivery of extra dose to those areas where it is required and/or the sparing of areas that are detrimentally sensitive to radiation. Figure 1.1 shows two immediate benefits: reducing normal tissue dose or increasing tumour dose while leaving the other dose unchanged will further increase the therapeutic ratio of the treatment. The radiobiological implications of possible errors with the delivery of IMRT to the patient will be discussed in depth later in this thesis and explained in terms of quantitative radiobiology.

The escalation of target dose and the reduction of dose to nearby healthy tissue both have the same effect of increasing dose gradients as there is a greater change in dose over the same distance. The necessity for appropriate margins is therefore critical.

**Defining treatment volumes**

Margins are used in radiotherapy to cover a range of scenarios to ensure that the target tumour is irradiated sufficiently during treatment (ICRU 1993, ICRU 1999). The gross tumour volume (GTV) is that defined by an oncologist as the observable tumour

when examining images of a patient, such as diagnostic X-ray or more commonly CT or magnetic resonance images. The GTV is usually expanded in three dimensions by the oncologist to a clinical target volume (CTV) to account for any sub-clinical spread of the tumour that cannot be seen on imaging and which will impair tumour control if not irradiated. The CTV may be clipped in some directions to account for impossible expansion, such as when a tumour is alongside a bony structure such as the skull or is very close to the patient's skin. The CTV is then grown in three dimensions to a planning target volume (PTV). The size of the CTV to PTV margin varies considerably depending on many factors: setup reproducibility of the patient on a day-to-day basis for multi-fraction treatments; inter-fraction movements of the target volume due to internal organ motion; intra-fraction movement of the target volume due to breathing, bladder/bowel filling or other patient motion. The CTV to PTV margin is not clipped in any direction as the uncertainties that it covers are not limited in any particular direction. However, variable margins may be used, such as with lung tumours, where a superior-inferior margin may be larger than the left-right margin to account for tumour movement during breathing. The PTV margin is conceived such that by planning a treatment to irradiate the PTV, treatment planners aim to achieve coverage of the CTV with 95% of the treatment dose for 95% of the time (van Herk *et al* 2000, McKenzie *et al* 2000). The size of the PTV margin can be calculated using the formula $M = 2.5\Sigma + a + b + \beta\left(\left(\sigma^2 + \sigma_p^2\right)^{1/2} - \sigma_p\right)$ (BIR 2003:Chapter 2), where $\Sigma$ is the total vector systematic error, $\sigma$ is the total vector random error, $a$ is the photon beam calculation error, $b$ is a vector representing breathing motion, $\sigma_p$ is the width of the beam penumbra and $\beta$ is a multiplier based on the number of overlapping beams. Typical values of $\beta$ range from 0.52 for in-plane margins for 6 co-planar non-opposed beams to 1.64 for the superior-inferior direction (assuming coplanar beams in the transverse plane). The formula can be approximated as $M = 2.5\Sigma + 0.7\sigma$ (van Herk *et*

*al* 2000) for a typical value of $\sigma_p$ and clinically significant values of $\sigma$ up to 7 mm or so. This shows that systematic errors have a much larger effect on the size of the PTV margin than do random errors. This is partly because random errors add in quadrature to the beam penumbra width. Reducing random errors therefore does not reduce the contribution to the overall margin size as much as reducing the systematic errors. Although the above formula details how errors should be added to achieve a PTV margin that gives 90% coverage of the CTV, the sizes of the errors still needs to be identified accurately.

It is clear therefore that any improvement in dose delivery that IMRT offers relies on appropriate margins and errors defined for each tumour, which must account for a wide range of variability. Work presented here will deal with patient setup reproducibility and the effects that this has on the PTV margin.

Intensity modulation is clearly a very complicated treatment modality and thus there is wide scope for uncertainties to be introduced into the doses the patient receives and the effects they will have on the tumour and healthy tissues within the patient. Work presented in this thesis will examine a number of aspects of the IMRT process, analysing the possible errors that can be introduced during the process, and putting forward suggestions and procedures that can be used to reduce those errors and improve the outcome of IMRT treatments for patients. The following aspects will be investigated:

Setup accuracy: Chapter 2 will focus on patient movement and setup accuracy within the radiotherapy treatment room, focussing on immobilisation devices..

Dose analysis methods and metrics: Chapter 3 evaluates different dose comparison and evaluation metrics and introduces new techniques for comparing doses.

Systematic errors: Chapter 4 will look at systematic errors in radiotherapy delivery that are accentuated by IMRT over 3D-CRT techniques, and how they affect the accuracy of IMRT in terms of the dose differences from a 'perfect' treatment.

Radiobiology and dose evaluation: Chapter 5 will present work based on a radiobiological grounding for choosing dose evaluation tolerances and investigate if there is equivalence between dose difference and distance to agreement.

Quantitative radiobiology: Chapter 6 will look at how the systematic errors analysed in Chapter 4 affect the radiobiological outcome of IMRT treatments, including calculations of patient outcome probabilities.

Drawing together the aspects previously highlighted, a final discussion will show where some errors may lie in IMRT treatments, how they can be reduced and how this improvement in accuracy can lead to better patient outcomes. A system that allows for errors to be incorporated and evaluated at the treatment planning stage will be developed, with a changing of focus in treatment planning from dosimetric accuracy to quantitative patient outcome.

# Chapter 2 – Patient movement and setup accuracy

## Introduction

As discussed in chapter 1, knowledge of patient movement and setup accuracy is of particular importance in the outcome of an IMRT treatment. Delivery of effective radiotherapy to a patient using external radiation beams requires that the treatment machine and the patient are accurately aligned, and reproducibly so if the treatment is divided into more than one fraction or if the treatment is based on pre-treatment imaging of the patient's anatomy. Treatment linacs and patients are free to move in many directions independently of each other – the gantry and couch can each rotate about the machine's isocentre with perpendicular axes and the collimator can rotate about an axis through the isocentre that is coincident with the radiation beam axis. The patient couch can also be moved laterally, longitudinally and vertically. For a multi-field radiotherapy treatment, the couch is moved laterally, longitudinally and vertically to position the patient before treatment of the first field. For subsequent fields, only isocentric movements are made (gantry, collimator and couch rotations), thus keeping the isocentre at a fixed point in the patient. Correctly aligning the patient in the linac's frame of reference and relative to the isocentre is crucial to ensure that the radiation beam is targeting the correct area of the patient. Immobilisation devices are frequently used to assist in achieving the same patient position for each treatment fraction (Waite and Filshie 1990, Hess *et al* 1995). Immobilisation devices take two forms – those that help to reduce patient movement and variability and those that also force an absolute patient position. Examples of the former are devices such as knee rests, which help to reduce hip roll for patients undergoing abdominal treatments, or ankle stocks to prevent leg abduction and associated movement of the hips. Patients using these devices may have external markers on their bodies, such as small indelible tattoos, to provide fixed

points to align with the treatment machine. The latter form of immobilisation device is usually a plastic cast individually moulded to the patient's shape. This is then fixed to the treatment couch so that the patient, when properly fitted into the cast, is positioned relative to the couch. The couch has a digital read-out of its position so the couch (and therefore patient) can be moved to the same place for every treatment fraction. Linac rooms have laser lines positioned to intersect at the isocentre, and marks are usually made on the exterior of the plastic cast to align with these lasers. There are therefore two methods of positioning the patient on the couch – either by using the digital read-out or aligning the marks to the lasers. Typically, the laser lines may be used on the first day of treatment and the couch read-outs are recorded and then used for subsequent fractions. A study has claimed (Greer 1998) that setting up to the couch height can reduce variations in patient position over setting up to markers, and so this technique may be more commonly used.

Immobilisation devices are most frequently fabricated for patients undergoing treatment to the head and neck area as this area typically requires better alignment between the external beams and internal structures due to the small size of target volumes and proximity of critical structures such as the brain stem and spinal cord.

A common form of head-and-neck immobilisation device is a mask made to fit over the patient's head while they lie supine on the couch, and the mask is fixed down to the couch, effectively fixing the patient's anatomy to the couch. Some mask systems extend further down the neck and across the shoulders, and systems vary in the number of fixation points used to fix the mask to the couch.

Two forms of immobilization mask are in common use: clear hard plastic shells made by vacuum-forming over a plaster cast of the patient, and thermoplastic material moulded directly over the patient. Both methods involve the patient lying on a mock-up treatment couch. The plaster-cast method begins by taking a negative impression of the

**Figure 2.1:** PETG shell mask shown with white quick-release poppers providing 11-point fixation to the couch top

patient using bandages and plaster of Paris which is left to set while on the patient. A positive bust is then created by filling this shell with plaster. The cast is sanded smooth, and then a transparent plastic is vacuum-formed over the cast. The mask thus formed is then drilled to create fixing points for the couch fixation system (see figure 2.1). Plastics commonly used are PETG (glycolised polyethylene terephthalate), PVC (polyvinyl chloride) and PMMA (polymethyl methacrylate). The thermoplastic mask method involves heating a sheet of opaque thermoplastic material to between 55 and 80 °C in a water bath or hot air stream. When it has become flexible and formable, it is placed directly over the patient and fixed down to points that match the treatment couch fixation points. The thermoplastic cools and sets while still on the patient. Because sheets of thermoplastic are opaque, they are supplied not as a solid sheet but as a meshed sheet. The holes are approximately 2 mm in diameter when the sheet is flat before forming, but stretch as the sheet is formed over the patient. This provides visibility of the patient during treatments, which helps to alleviate patient anxiety in the otherwise opaque mask, but also allows the treatment radiographers to see if the mask is too tight or too loose due to changes in patient shape and size during their course of treatment. Some thermoplastic sheets are provided with solid bars of material in the mesh sheet to provide extra stability of the formed mask. Such bars may be placed in the forehead or nose region of the mask, or across the shoulders (see figure 2.2 over the page)

No patient mask can be made as a perfect tight fit, as patients need not only accurate, but also comfortable treatment setup. Therefore there is always a tolerance of movement associated with any immobilisation device. The task of the manufacturers is to improve this error while retaining patient comfort.

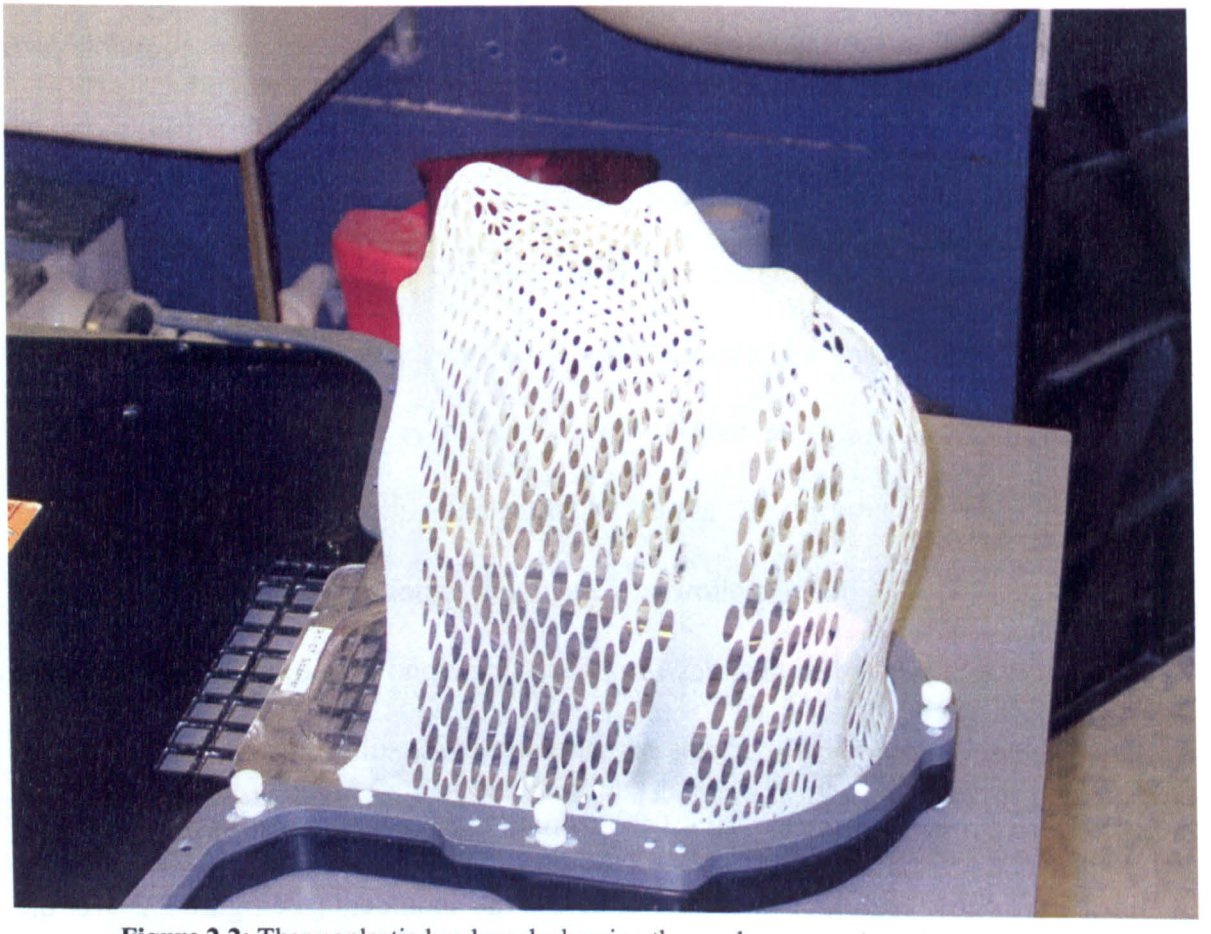


**Figure 2.2:** Thermoplastic head mask showing the mesh construction of the mask with strengthening bars running across the forehead and nose areas

There have been many studies to evaluate the positioning reproducibility of patient mask immobilisation devices (Hess *et al* 1995, Weltens *et al* 1995, Gilbeau *et al* 2001, Humphreys *et al* 2005). However, very few have evaluated more than one or two forms of mask simultaneously. Many of the studies use portal imaging of patient treatments to assess day-to-day variability of setup. While this provides retrospective data on how patients have been setting up, it provides no theoretical framework for establishing maximum limits on tolerances, and patients can be much more variable in their daily setup than well-informed volunteers. The measuring of patient setup also typically involves using portal images compared against gold-standard imaging such as CT data or digitally reconstructed radiographs (DRRs) generated from CT data. Not only does this require registering the portal image to the DRR, which may involve errors of over 1.5 mm (Hawkes 1998, Jobse *et al* 2003, Gilhuijs *et al* 1996), but the pixel size of portal images ranges from 0.25 mm to over 1mm, with a typical size of 0.5 mm (Langmack 2001, Boyer *et al* 1992, Munro 1995). Thus the accuracy of results from such evaluations may be poor. Patients may also change size due to such factors as steroid intake associated with their treatment or reaction to radiation – both swelling of the irradiated area and shrinkage as the tumour recedes. These changes can affect how well a patient fits their immobilisation mask as their treatment progresses.

Other forms of head-and-neck immobilisation are used, such as mouthbite or mouthplate systems. These register bony anatomy of the mouth (either the palate or the upper jaw) to the couch system to identify patient position. Mouthbite systems use a custom-moulded teeth impression that the patient bites into every fraction – the mouth bite is then secured to the couch. Mouthplate systems use a custom-formed plate which is vacuum-attached to the top of the patient's mouth, and the plate is then fixed to the couch.

The issue of patient immobilisation is particularly important for IMRT as the dose distributions created inside the patient are more complex than those for 3D-CRT, typically 'wrapping' dose around critical structures. Large dose gradients near to radiosensitive organs means that even small setup errors can lead to excessive doses being delivered to those organs. The MSF technique for IMRT delivery also takes longer than 3D-CRT, with many more segments being delivered. Intra-fraction motion of a patient is therefore more likely and so effective immobilisation is critical.

The work presented here aims to evaluate the setup reproducibility of several head-and-neck immobilisation masks from several different manufacturers. The results will be compared against various other systems published in the literature. This setup study is performed on a healthy volunteer who can report feelings inside the mask in an articulate way and can also respond to instructions when inside the masks. The use of a healthy volunteer who is not undergoing treatment is important for maintaining accuracy, as many patients, used in other studies of immobilisation, undergo severe weight loss during the course of their treatment (Beaver 2001).

## Materials and methods

### 1) The subject

The subject is a healthy male, of slim build, aged 25, with short hair. The weight of the subject varied from 64.8 kg to 66.8 kg over the course of the study. The possibility of movement due to compression of the hair was therefore small, and the effects of changing anatomical shape were minimal. Only one subject was used to reduce costs of material, as 10 different immobilization mask methods were used and many of the systems used were trials from manufacturers. The study was carried out over several months as different materials became available.

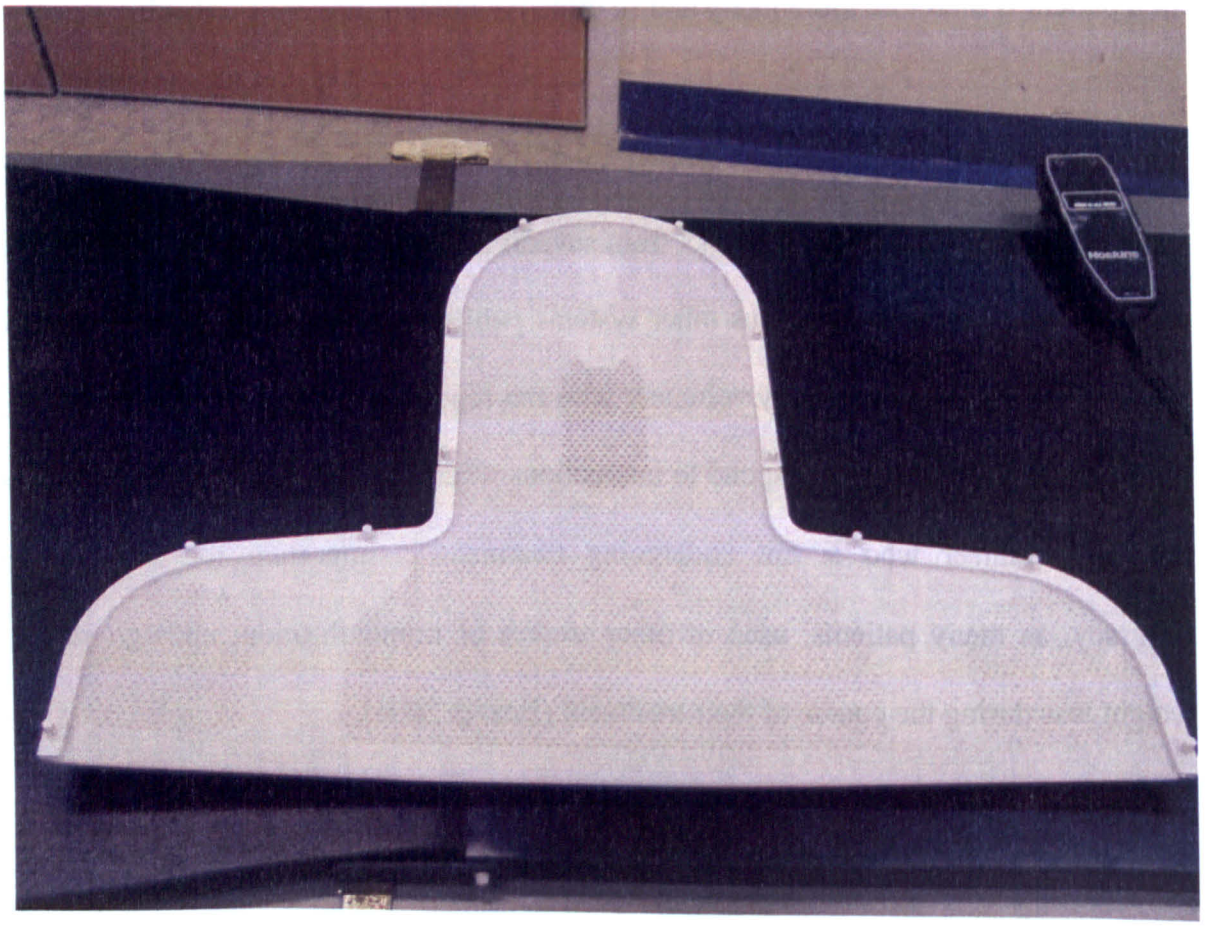


**Figure 2.3:** Thermoplastic mask material before forming over a patient. This image clearly shows the mesh structure of the material, with extra holes near the nose and mouth area, as well as strengthening bars of solid material across the brow/forehead and shoulders.

**2) The masks**

Two plaster-cast PETG shells were made:

1) Impression was taken with the subject's shoulders relaxed.

2) Impression was taken with the shoulders pushed towards the feet using Fixator[3] detachable pads on the table top.

Eight thermoplastic masks were made, of varying material and number of fixation points:

3) Orfit[1] 5-point mask covering the head, neck and shoulder tops;

4) Addisilk[2] 3-point mask covering the head only;

5) Imotek 5-point mask covering head, neck and shoulder tops;

6) Imotek 4-point mask covering head, neck and shoulder tops, but not the top of the head or the forehead;

7) Medtec S-Type[3] 5-point frame, covering only the head. The mesh had 4 strengthening bars – starting from just above the bridge of the nose, the bars run to the left and right of the patient and at 60° and 120° angles towards the top of the head;

8) S-Type 5-point frame covering the head only. The mesh has 3 strengthening bars, running left, right and superiorly from the top of the bridge of the nose;

9) Aquaplast[4] 9-point frame covering the head, neck and shoulders;

10) S-Type 9-point frame covering the head, neck and shoulders. The mesh used 3 strengthening bars across the forehead plus wider bars across each shoulder (figure 2.3).

All 10 setups used a hard plastic headrest to extend the neck and tilt the chin back. This is a common technique used at Addenbrooke's for head and neck patients as

[1] Orfit Industries, Wijnegen, Belgium
[2] Oncology Imaging Syetems, Uckfield, UK
[3] Medtec, Orange City, Indiana
[4] WFR/Aquaplast Corp, Wyckoff, New Jersey

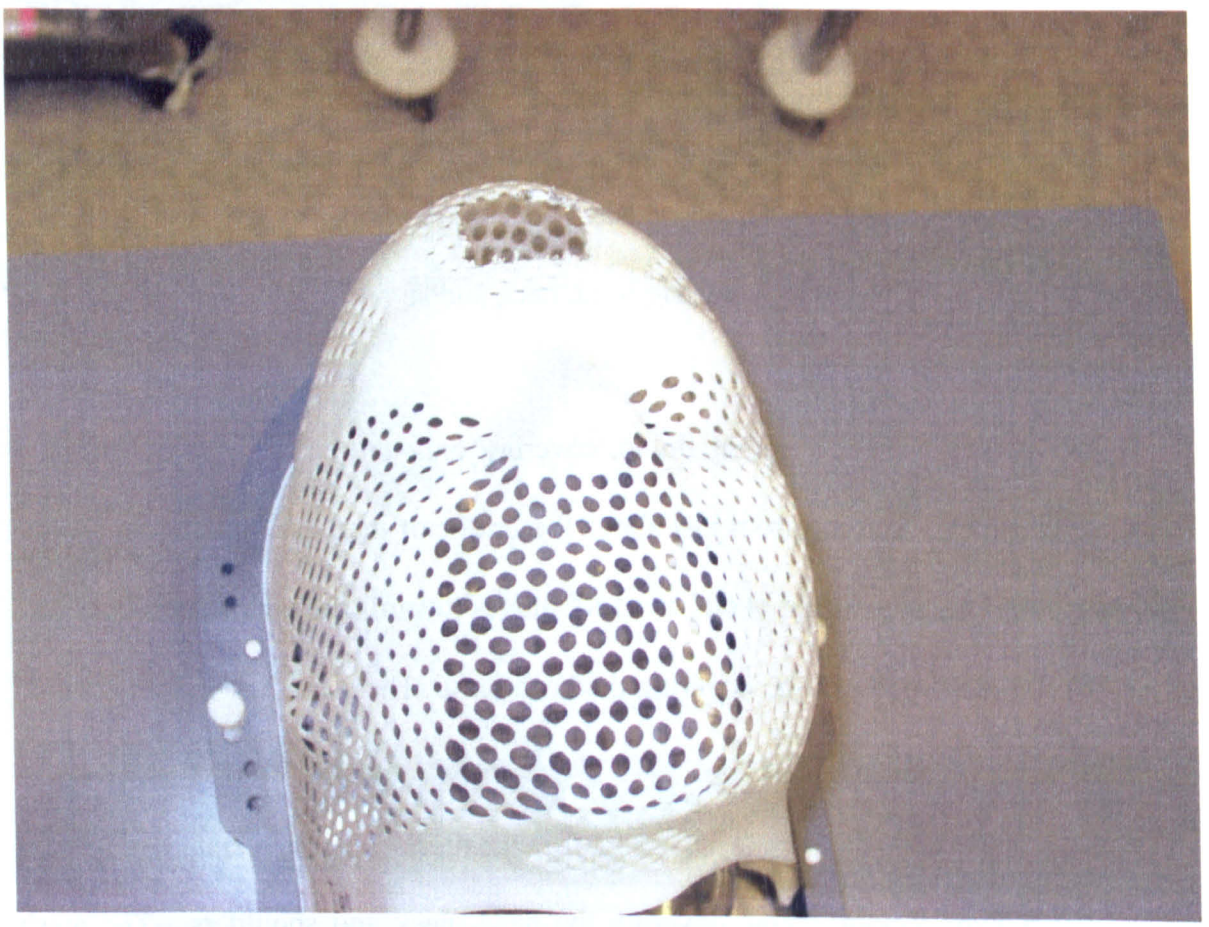


**Figure 2.4:** Thermoplastic head mask showing the cut-out used for the fiducial marker. The strengthening bars of solid material can be seen running across the brow and either side of the cut-out.

it moves the jaw away from the neck to spare it irradiation and also allows access to the neck area with electron applicators. The headrest used was a selected from a set of manufacturer-supplied ones, and chosen to provide the best anatomical fit to the patient's neck.

All of the shells and masks were made by the same mould room staff team, who have several years of experience using both thermoplastic and plaster-cast methods.

### 3) The measurements

Measurement of the setup reproducibility was performed using a Minolta[1] Vivid VI-700 non-contact 3D digitizer. This system uses a scanning laser line and camera to record a three-dimensional point model of an object in its field of view. The system saves data files to a memory card which are then read into the manufacturer-supplied software on a PC. The software exports all of the measured points in the original file as an ASCII point-cloud file, consisting of the 3D Cartesian co-ordinates of each point. This file was read into the MATLAB[2] programming environment using a script created by the author for this project. The (x,y,z) co-ordinate triples were displayed and measured interactively using the display routines within the MATLAB software.

Measurements were made by imaging the subject on a linac treatment couch with a fiducial marker attached to his forehead, the marker protruding past the mask through a 3×3 $cm^2$ window cut in the mask (see figure 2.4). Multiple sections larger than this are frequently removed from patient masks to reduce build-up of radiation at the skin, so it was not anticipated that a single window would cause structural weakening of the masks. As the coordinate system of the digitizer is relative to its own position, a fiducial marker was also fixed to the treatment couch in the field of view of

[1] KonicaMinolta, Mahwah, New Jersey
[2] The Mathworks, Natick, Massachusetts

the camera. All measurements of the subject's position were taken from the tip of this marker, thus showing the position of the subject relative to the treatment couch.

The accuracy of the laser scanning camera was established without the subject present by repeated images of fixed objects and by moving one fiducial marker by known distances.

For the measurements of setup, the subject was asked to move a maximum comfortable distance in each of the superior, inferior, left and right directions, and a scan was taken at each location. Anterior and posterior movements were not measured for this study as it was felt that movements in this direction are more affected by how the patient lies in the mask in terms of nervous tension in the neck and shoulders than how well the mask fits and prevents movement. Although there is likely to be ant-post movement of patients in immobilisation masks, as shown by previous studies, the magnitude of the movements are harder to quantify with a single subject.

Two baseline scans were taken of a 'central' position at the beginning and end of the measurement sequence with the subject asked to position himself in what felt to be the most central comfortable position. The set of 6 measurements were then repeated after the subject had been removed from the mask, allowed to walk around the room and then refitted into the same mask. The baseline scans were performed to ensure that the co-ordinate system of the camera was constant, and also to check the constancy of the central position of the subject. Movements in each of the 4 directions will be quoted relative to the average position of the baseline scans.

**Results**

The data in the camera's data file is given with 0.01 mm precision in each of the three Cartesian axes. It is not known if these values are obtained by truncation or

| Mask | Movements, mm | | | | Baseline Diff. (3D) |
|---|---|---|---|---|---|
| | Right | Left | Sup. | Inf. | |
| Imotek 4-point | 4.75 | 5.25 | 5.0 | 1.75 | 0.3 |
| Imotek 5-point | 3.75 | 4.75 | 3.75 | 2.0 | 0.4 |
| Addisilk 3-point | 4.5 | 3.75 | 4.0 | 2.75 | 0.3 |
| Orfit 5-point | 4.25 | 3.5 | 3.75 | 3.5 | 0.4 |
| Medtec 3-bar | 4.25 | 3.75 | 3.75 | 3.75 | 0.3 |
| Medtec 4-bar | 4.0 | 3.25 | 2.75 | 3.5 | 0.4 |
| Medtec 9-point | 3.25 | 3.0 | 2.5 | 1.5 | 0.2 |
| Aquaplast 9-point | 3.25 | 3.5 | 3.5 | 3.0 | 0.4 |
| PETG shell | 5.0 | 4.75 | 5.25 | 4.0 | 0.5 |
| PETG shell + shoulder retractors | 4.5 | 4.25 | 3.25 | 3.25 | 0.5 |

**Table 2.I:** Results of the measurements on the mask types listed. The movements (in mm) are the average of two sets of measurements, taken as distances from the average of the baseline images. The difference between the baseline images is also shown as a 3D vector distance.

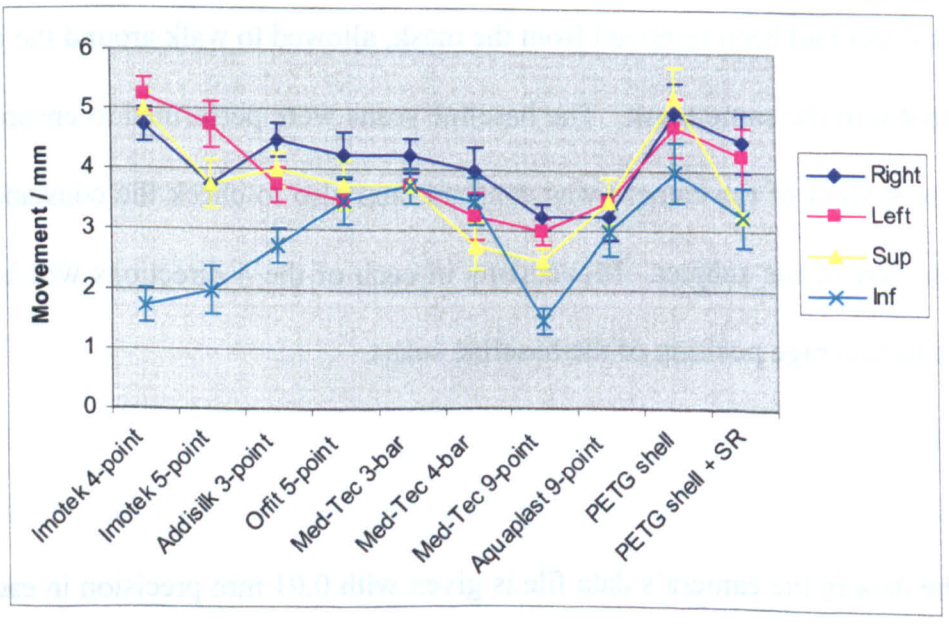



**Figure 2.5:** Maximum movement of the subject within each of the masks shown, in each of the right, left, inferior and superior directions. The error bars show the standard deviation of the baseline scans.

rounding, but the precision of the values is much greater than the anticipated accuracy of measurement, and the rounding error is therefore ignored. The known movements of the objects used in the assessment of the measurement accuracy could be measured only to the nearest 0.25 mm. The measurements of the locations of these objects from the digitized images were within 0.25 mm of the expected positions, and therefore the precision of the 3D coordinates exported from the digitizer is taken to be ±0.25 mm (which equates to a standard deviation of 0.14 mm – see below).

The results of the subject's measured movements are presented in table 2.I, showing movement in each of the 4 directions (right, left, superior, inferior). The standard deviation of the four baseline scans is shown. The results show the average of the two sets of readings taken, rounded to the nearest 0.25 mm and are presented graphically in figure 2.5.

Table 2.II (over page) shows the results of setup studies that have been published in the literature for comparison to the results obtained here.

Harrison and McKenzie (BIR 2003:Appendix 2b) show that the standard deviation of a top-hat function of half-width $a$ is equal to $a\sqrt{3}$. Table 2.Ia (over page) therefore shows the measured values from table 2.I divided by $\sqrt{3}$ to give a standard deviation. This then presents the results of this study in the same format as those from other studies in the literature (table 2.II), which are all quoted to 1 SD.

**Discussion**

The results from measurements taken for this study show that all systems gave positioning of better than 5.25 mm from the central location, with the standard deviation calculated as less than 2.9 mm for all systems. This agrees well with previously published studies presented in table 2.II, which show a maximum standard deviation of

| Mask | RL | SI | AP | 3D | Reference |
|---|---|---|---|---|---|
| Orfit 5-point | - | 2.1 | 2.1 | - | Weltens *et al* 1995 |
| PETG shell | - | 2.1 | 2.1 | - | Weltens *et al* 1995 |
| Posifix 3-point[1] | - | - | - | 3.1 | Gilbeau *et al* 2001 |
| Posifix 4-point | - | - | - | 2.4 | Gilbeau *et al* 2001 |
| Posifix 5-point | - | - | - | 2.4 | Gilbeau *et al* 2001 |
| Aquaplast | 1.3 | 1.0 | 1.2 | - | Tsai *et al* 1999 |
| Orfit 3-point | - | - | - | 3.05 | Sweeney *et al* 1998 |
| Mouthplate system | - | - | - | 1.02 | Sweeney *et al* 1998 |
| Mouthbite | 3.1 | 2.5 | 2.7 | 3.1 | Willner *et al* 1997 |
| PVC shell | 0.6 | 0.5 | - | - | De Boer *et al* 2001 |
| Orfit Raycast (3 point) | 0.74 | 0.93 | 0.75 | 1.59 | Fuss *et al* 2004 |
| UON precise-fit[2] | 1.2 | 1.1 | 0.6 | - | Prisciandaro *et al* 2004 |
| Medtec S-Type 9-point | 0.3 | 1.1 | 0.8 | - | Prisciandaro *et al* 2004 |

**Table 2.II:** Results of mask setup studies previously published showing right-left (RL), superior-inferior (SI), anterior-posterior (AP) and 3D displacements. Movements are quoted as the standard deviation about the mean.

| Mask | Movements, mm | | | | Baseline Diff. (3D) |
|---|---|---|---|---|---|
| | Right | Left | Sup. | Inf. | |
| Imotek 4-point | 2.6 | 3.0 | 2.9 | 1.0 | 0.3 |
| Imotek 5-point | 2.2 | 2.7 | 2.2 | 1.2 | 0.4 |
| Addisilk 3-point | 2.6 | 2.2 | 2.3 | 1.6 | 0.3 |
| Orfit 5-point | 2.5 | 2.0 | 2.2 | 2.0 | 0.4 |
| Medtec 3-bar | 2.5 | 2.2 | 2.2 | 2.2 | 0.3 |
| Medtec 4-bar | 2.3 | 1.9 | 1.6 | 2.0 | 0.4 |
| Medtec 9-point | 1.9 | 1.7 | 1.4 | 0.9 | 0.2 |
| Aquaplast 9-point | 1.9 | 2.0 | 2.0 | 1.7 | 0.4 |
| PETG shell | 2.9 | 2.7 | 3.0 | 2.3 | 0.5 |
| PETG shell + shoulder retractors | 2.6 | 2.5 | 1.9 | 1.9 | 0.5 |

**Table 2.Ia:** Figures from table 2.I, divided by $\sqrt{3}$ to convert the maximum displacements to standard deviation, after Harrison and McKenzie. Figures are shown to 1 decimal place. The baseline difference column is unaltered.

[1] Gammex rmi, Middleton, Wisconsin
[2] Nuclear Associates, New York

3.1 mm. Work by Harrison and McKenzie shows that the standard deviation of a top-hat function can be calculated as $a/\sqrt{3}$ for a half-width of $a$. Using this method to determine the equivalent standard deviation of the measurements taken here implicitly assumes that the variability of patients within immobilisation masks is a top-hat function. For a single subject, as used in this study, the distribution of positions within a mask over a number of measurements would clearly not be a normal, Gaussian distribution, as there is a definite cut-off point beyond which there is no more possible movement. There is a maximum possible displacement of the subject, which is measurable, and there is no logical reason why the patient cannot assume any position within that maximum at each refitting of the mask. The probability distribution of a single patient is therefore likely to be a top-hat distribution. By taking a standard deviation of that top-hat, not only is it possible to compare results with other studies that involve many patients and quote standard deviations as variability results, but it may be possible to infer the distribution of a population from a single subject. Therefore while the results of this study, when converted to standard deviations, are not exactly comparable with other statistical methods used, important conclusions can be drawn from this study and comparisons can be made. Henceforth reference will be made to table 2.Ia when discussing the results of this study.

The measurements show a wide variation in the immobilisation of the subject, from less than 1 mm for the inferior movement with some systems, to as much as 3 mm for others. Figure 2.5 shows that with the exception of the Medtec 5-point system with 4 strengthening bars, the inferior movement was the smallest of the four directions. The largest movements were recorded mainly for the left or right directions. When examined without the subject, the thermoplastic masks showed greatest flexibility in the left and right directions, with excellent rigidity in the superior-inferior directions. The left-right flexibility accounts for most of the measured movement, as in general the

masks fitted well to the sides of the head. When a thermoplastic mask is formed, the greatest stretching of the mesh occurs down the sides of the head, as can be seen in figure 2.2, where the gradient of the patient's head is the greatest. This inevitably leads to less structural strength and rigidity in these areas, allowing the mask to flex from left to right. Although the PETG shells were far more rigid in all directions when fixed to the couch top, they still allowed significant left-right movement of the subject. This is most likely due to the manufacturing process creating shells that are too large. The negative cast of the patient, formed from plaster bandages, is only a few layers of bandage thick and is not very strong. This is in order to reduce the material costs of bandages used and the time spent by the patient covered in the cast. When the positive bust is made it is likely that the weight of plaster in the negative bows it out and leads to a bust which is larger than the patient. The negative plaster shell is supported with bean bags during the plaster filling process, but the results of this study imply that there is still error in the process as the PETG shell should provide a snug fit.

As expected, the Imotek 4-point fixation system showed a large movement in the superior direction, due to the lack of any material at the superior end of the head. The lack of any support at 90° to the left-right direction also gave this mask more flexibility that way, reflected in the slightly higher left and right measured movements. There is little bony anatomy to prevent movement in the superior direction except for the bridge of the nose and the shoulders, and a less-than-perfect fit of the mask will allow significant superior movement, as is shown. In contrast, the inferior movements for most masks was good, primarily due to good contact between the mask and vertical surfaces of the subject such as below the chin and the base of the nose, both of which prevent significant inferior motion. The Imotek masks, in particular, have a cut out in the material for the nose, such that it protrudes through. There is little stretching of the material around the nose area, and there is therefore a hard edge to the material around

the nose cut-out. This hard edge made it uncomfortable for the subject to move in the inferior direction, and this is reflected in the measurements. The Medtec 9-point system had a small hard ridge at the base of the nose which also made inferior movement uncomfortable.

PETG or PVC shells are often regarded as a gold standard for immobilisation, because the process is more established than the use of thermoplastics and because of the rigidity of the shell. It is interesting to note that the measurements recorded here show that one of the worst immobilisations in this study came from the PETG shell. In addition to the bowing of the plaster negative detailed above, this may be due to patient/subject tension. The process of casting involves covering the entire head, mouth, eyes and ears with plaster bandages, leaving only nose breathing holes free. This process naturally leads to some tension in the patient or subject, even though for this study he was well-informed and volunteered for the process. At the time of the measurements, there was a noticeable gap around the shoulder areas, implying that there was a significant difference in relaxation between casting and refitting. In contrast, the thermoplastic method uses a mesh which leaves access to the eyes, ears, nose and mouth throughout the procedure. With many of the thermoplastic masks used, there was a much smaller gap between shoulders and the material (for those masks which covered the shoulders).

With the PETG shell, the addition of shoulder retractors improved the movements in all four directions, especially to the superior. With the shoulders pulled towards the feet there was less scope for hunching of the shoulders and therefore a better fit of the shell later. The retraction of the shoulders also extended the neck, which made superior movements difficult. Without shoulder retraction the head can move superiorly without moving the shoulders by extension of the neck.

As with previous studies, the results presented here do not show a clear trend towards better immobilisation for more fixation points, although there is better immobilisation as the fixation improves for a particular material. For example, the Addisilk 3-point fixing system is better than the Imotek 4-point system (except for inferior movements), but the Medtec 9-point system is better than the 5-point 4-bar system, which in turn is better than the 5-point 3-bar system.

In general, the movements measured agree well with previous studies using similar materials. For example, Weltens *et al* report a 2.1 mm standard deviation for the Orfit 5-point system, compared to 2.0-2.5 mm measured at Addenbrooke's. Prisciandaro *et al* report 1.1 mm for superior-inferior movements, compared to 0.9-1.4 mm here. There are also notable differences in the results, for example with Aquaplast, with 1.7-2.0 mm movements measured comparing to 1.0-1.3 mm reported by Tsai *et al*. The PETG shell, measured between 1.9-2.6 mm (with shoulder retractors) compares well with Weltens *et al* 2.1 mm, but not well with De Boer *et al* measurements of 0.6 mm for a PVC shell. Some of the differences may come from the measurement process used. Many of the previous studies have used portal imaging of patients to assess setup reproducibility, as compared to measuring maximal values in this study. Although this study shows that maximum movements within a PETG shell may reach 5 mm, patients may reposition more consistently for several reasons. The head rest used for the subject here, for example, was relatively flat across the occipital portion of the head. If a more deeply bowled headrest is used, a patient is more likely to attain a central position rather than significantly to the left or right. Radiographers also play a key role. When setting up patients for their treatment, radiographers attempt to get the most comfortable position for a patient, and may ask the patient to move slightly if they can see that the patient is pressing against the mask. That contrasts with this study, in which a

deliberate attempt was made to find the range of movement *possible* rather than the movement that is *likely*.

This study is unique among those previously published in terms of the number of different materials tested, with previous studies testing no more than 3 different masks. Although this study does not provide precise population information on how a range of patients will reposition on a daily basis, it does provide important information. Because all of the masks were made by the same mould room staff team, and all used the same subject, the relative differences between the masks show which materials are the best to use in the setting of Addenbrooke's hospital. It may be that certain staff are better with certain materials, for example, or that some thermoplastics require more staff to hold the material in place while it sets on the patient. This study has also allowed the use of a volunteer who could articulate what each mask is like in terms of the fit and the flexibility within it, whereas many patients, only having experience of one mask, cannot relate such information. Additionally, this study used one subject whose weight and size did not change significantly throughout the whole process. Many patients experience shape change due to steroid intake, swelling from radiosensitivity or weight loss during the course of their treatment. Thus previous studies have shown how patients reposition over a course of treatment when these effects are taken into account, whereas this study shows the immobilisation capabilities of the mask only. Some shape change can be compensated for by adding or removing sheets of plastic placed beneath the headrest. It is important to be able to separate patient shape change and immobilisation as this study has done.

This study has shown that a non-invasive method is possible for measuring the position of volunteers or patients in immobilisation masks. Laser-based systems such as the digitizer used here, can provide safe, repeatable measurements that could be used on

a large patient set for measurement of population statistics, as have been reported previously. The benefit of this would be that the patients would require no extra portal imaging or radiation dose and scans could be repeated without health and safety or ethical issues. Measurement of object positions using a laser digitizer are also less prone to errors than the registration of portal images to DRR imaging, and is likely to be more precise and more accurate.

This study has measured the setup accuracy of several different immobilisation systems. The first step on the way to improving accuracy is to measure that achievable at present. The results presented have shown that the current PETG shell method that is considered as the best immobilisation system is actually bettered by several thermoplastic systems and that the Medtec 9-point system is the best system to use in the current environment at Addenbrooke's. By measuring and understanding the setup accuracies of patients, moves can be made to alter treatment margins in use for both conformal and intensity modulated radiotherapy. By moving patients over to the most accurate system available, the setup error is reduced. Treatment margins are calculated as the addition in quadrature of several potential errors. Therefore, provided that the setup error is not small compared to the other potential errors, reducing it will allow for the reduction of treatment margins. This will still ensure the coverage of the CTV during radiotherapy, but will allow the reduction of the irradiated volume, which will be a reduction of normal tissue irradiation. This will lead to improved patient outcome either through a reduction in normal tissue complication or by allowing dose escalation of the target volume. Both of these would have the effect of increasing the probability of uncomplicated control (see chapter 1).

This study used a single subject, so cannot show the appropriate treatment margins to use for a population of patients. It can show the best immobilisation system

to use, and from there future work needs to focus on measuring the movements of patients in order to derive treatment planning margins. Having shown that the PETG shell manufacturing method may lead to creating a mask that is too large, further investigation could confirm this and help to reduce it. A hand-held laser-based digitizer, FastScan[1], can be used to create 3D models of a subject which can then be sent directly to manufacturing to create a positive bust by milling from plaster blocks. By creating such a positive bust, the traditional shell-making process of plaster bandages and back-filling could be used to make a PETG shell for the bust. Any problems with subsequent fitting of the shell would highlight inaccuracies in the shell-making process. Ultimately, the FastScan system can eliminate the need for the plaster bust stages of the process, although at present milling machines large enough to create head-and-neck busts of patients are prohibitively expensive for most radiotherapy centres.

**Conclusions**

A method has been described to establish the setup accuracy of immobilisation devices without the use of ionising radiation. A laser-scanning camera can assess the movement potential for a subject inside an immobilisation device with good accuracy and precision. Because ionising radiation is not used as part of the imaging process, the same subject can be used for several mask systems to assess which is the best one. By selecting the best masks and with knowledge of the likely movements, improvements can be made in patient setup and in treatment planning, which will lead to better accuracy in radiotherapy treatments and improvements in patient outcome.

[1] Polhemus, Vermont, USA

# Chapter 3 – Dose analysis methods and the introduction of Kappa, a new dose evaluation metric

## Introduction

At some stage during the planning process of a radiotherapy treatment, quality assurance and checking of the calculated delivered dose is carried out. Dose data are compared and the comparison must lie within certain tolerances in order for the treatment to be permitted. The method of comparison of such data depends on the source of the data, as doses can be measured and calculated in 0, 1, 2 or 3-dimensional format (point-dose, linear profile, plane profile or full 3D matrix respectively). The comparison of a single point-dose, for example, is a trivial matter, whereas the comparison of calculated and measured 3-dimensional dose distributions is a less easy task, both in analysis of the data and visualisation of it.

The IMRT planning and delivery process is more complex than traditional 3D-CRT and by its nature introduces errors and accentuates those associated with 3D-CRT (see chapter 5 for more discussion). These errors will manifest themselves in the dose distributions created and therefore there is a need for a rigorous, yet effective method to compare ideal, calculated dose distributions and the inevitably non-ideal dose distributions measured during QA work. The same dose analysis methods are also vital in order to assess potential errors and problems associated with the IMRT process.

The quantitative and effective analysis of dose data across 2D planes and 3D matrices requires computerized techniques. The work presented here aims to review the dose analysis methods commonly in use as reported in the literature and identify the strengths and weaknesses with each in terms of their effectiveness of comparing ideal and non-ideal dose distributions. A routine will be presented for optimizing the speed of the calculation algorithm for the most common, and slowest, metric. Further

refinement to the metric will be presented as an algorithm for calculating that metric in the most accurate way. Finally, a new dose evaluation metric will be presented which will aim to circumvent some of the problems associated with the current methods in use.

For the discussion of dose distributions and comparisons, the terms 'ideal distribution' and 'non-ideal distribution' will be used. These are analogous to the 'reference' and 'evaluation' distributions as suggested by Low and Dempsey (2003). The ideal dose distribution is the dose distribution as planned or calculated by a treatment planning system, and the non-ideal distribution is data measured from a treatment machine or generated by a secondary treatment planning or calculation system. The ideal dose distribution is always the 'gold standard' against which the non-ideal distribution is compared, and in the case of patient treatments, the ideal distribution is that recorded as the dose the patient received. The terms 'ideal' and 'non-ideal' distributions will be used to refer to 2D planes and 3D matrices of data.

## Dose evaluation metrics – a review

### Dose difference

The simplest dose comparison tool for any dimension of data is a dose difference. The dose at every point in the ideal dose distribution is subtracted from the dose at an equivalent point in the non-ideal distribution. The size and shape of the matrix of data produced is of the same order as the original distributions. If the ideal and non-ideal dose distribution data have different data point spacing then one set of data will require interpolation to match the other. Although interpolation can be slow it only has to be performed once and the subsequent dose difference analysis is very fast to compute. The resulting data is easy to analyse visually if presented as a colour wash with different colours corresponding to different dose differences. In practice,

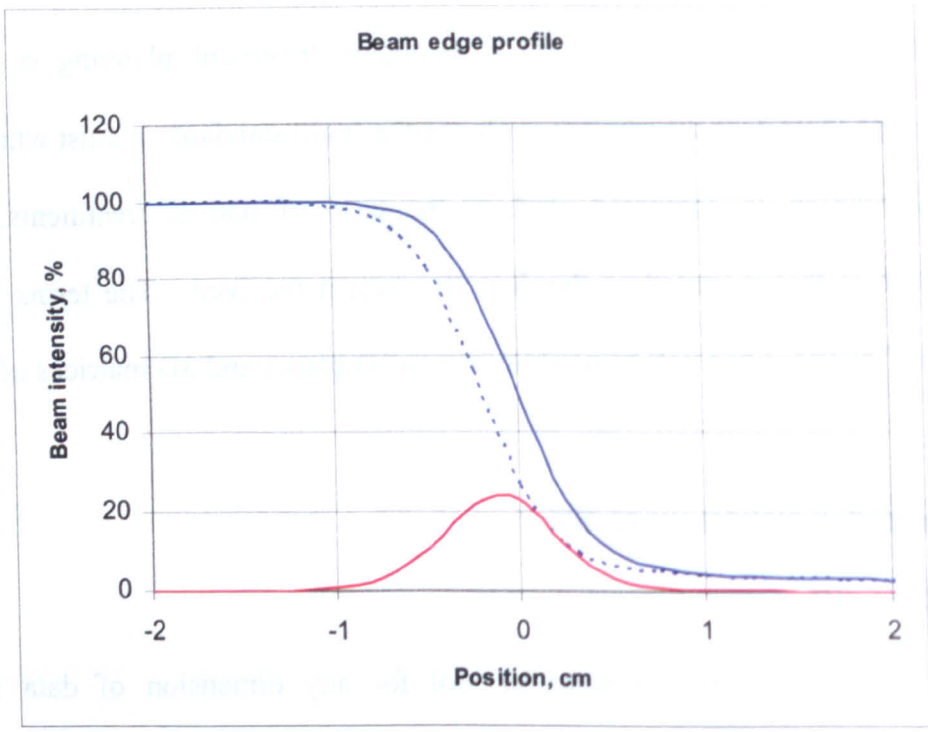



**Figure 3.1** A beam edge penumbra profile measured on a Siemens linear accelerator. The solid blue line shows a normal penumbra, with the 50% field edge defined at the 0 cm position. The dotted blue line shows the same profile with a 2mm lateral shift to the left. The red line at the bottom shows the difference between the two penumbrae, with a maximum difference of 24% of the beam intensity

visualisation of 3-dimensional arrays of data is difficult, so data of this dimension will usually be shown as a series of 2-dimensional planes of colour wash.

Although dose difference data is easy to visualise, it is difficult to obtain an overall quantitative measure of the congruence of two dose distributions. If, for example, the distributions are 3-dimensional, covering a 20×20×20 $cm^3$ volume at a pixel spacing of 2 mm, there are over 1 million data points ($101^3$) in the analysis. A simple mean and standard deviation of the dose difference matrix can be useful, although collapsing a million data points to two values in such a way will tend to lose important data. For example, a 1 $cm^3$ volume could receive a very large overdose error without affecting the mean and standard deviation by a significantly detectable amount. If this volume corresponds to a sensitive organ in the patient treatment (such as a section of the spinal cord), then the mean and standard deviation of the dose difference would clearly give poor analysis of the data. A report of the range of the data in the matrix, showing the maximum and minimum dose differences, is usually unhelpful as single erroneous data points will skew the range significantly. These erroneous points may come from specks of dust on scanned films or a spike of data from a single detector in a solid-state array. While a visual inspection of a colourwash plane of dose difference data can easily identify single data points that are clearly incorrect, automated methods for removing them are prone to introducing errors. Techniques for automatic removal may rely on removing spikes below a certain size, or the range could be reported as the range of the central $n$% of the data values. Both methods require some decision of a threshold, and important data may be removed or not removed near the thresholds.

Additionally, a small spatial shift of one dose distribution relative to the other may produce very large dose differences in areas of steep dose gradient, as demonstrated for a single beam penumbra in figure 3.1. IMRT treatments, especially,

tend to have many steep gradients within the dose distribution, so a small spatial shift within acceptable margins may give an incorrectly poor analysis of the congruence of the ideal and non-ideal dose distributions. Thus simple dose difference is a limited tool for the quantitative analysis of dose distributions if used on its own.

**Isodoses and distance-to-agreement**

Physicists and oncologists may wish to view the congruence between isodose lines in the ideal and non-ideal dose distributions to evaluate them. Overlaying two isodose plots can show an overall agreement between two distributions and the distance that one distribution may be shifted relative to the other. In order to provide quantitative measurement of the isodose separation and circumvent the problems of dose difference analysis of steep dose gradients and small spatial shifts, the concept of distance-to-agreement (DTA) can be used (Hogstrom *et al* 1984, Dahlin *et al* 1983, ICRU 1987, Van Dyk *et al* 1993). The DTA shows, at every point across one dose distribution, the physical distance from that point to the nearest point in a second distribution that has the same dose value. The measure of DTA is clearly linked to the concepts of treatment margins – a tolerable DTA can be included in the CTV to PTV margin to account for accuracy of delivery of a dose distribution to the target volume.

In high dose gradient areas, the DTA measure is very useful, as it shows how physically close the ideal and non-ideal distributions are to matching. However, like dose difference, DTA produces a matrix of data that is the same size and dimension as the original dose distributions, with many hundreds of thousands of points for 3-dimensional data. DTA exhibits similar problems to the dose difference method although they manifest themselves in low dose gradient areas rather than high gradient areas; in a low dose gradient area, a small change in dose will correspond to a large change in distance. For example, if the dose gradient is only 0.05 cGy / mm, a 1 cGy

dose difference would give a DTA of 20 mm. Since the aim of treatment planning is to achieve a uniform dose across a target volume and such low dose gradients are common, DTA as a dose checking measure is less useful than dose difference in these areas.

As with dose difference, DTA matrices may be analysed by viewing colour washes, or by calculating the mean and standard deviation although again these analyses tend to smooth out problem areas in the analysis matrix and thus are not very useful. As such, although dose difference and DTA both provide quantitative data, they tend only to be used as qualitative measures of dose distribution congruence for all except the most simple cases such as point measurement.

A method to provide more quantitative answers from DTA and dose difference result matrices is to place tolerances upon them. The prescribing oncologist will tolerate certain errors in the delivery of the treatment plan, for example a difference of 3% of maximum dose or 3 mm DTA. Thus DTA and dose difference analysis matrices can be thresholded at these levels to produce binary displays of those areas that pass tolerance (0) and those areas that fail (1). Although showing immediately those areas of a dose distribution that pass or fail, rather than a confusing colour wash of all available data, this method can only provide partly quantitative data to the analysing physicist. The problem arises when areas are very close to tolerance. The oncologist may accept a non-ideal dose distribution where the whole is well within tolerance except for a few small areas that are just out of tolerance. On a binary display, this would be shown as less acceptable than a whole matrix of data which is just within tolerance and which might not be acceptable to the oncologist if presented as a colour wash. Additional to reporting a binary answer is the problem that analysis of either DTA or dose difference on their own proves inadequate for the whole range of dose gradients present in a

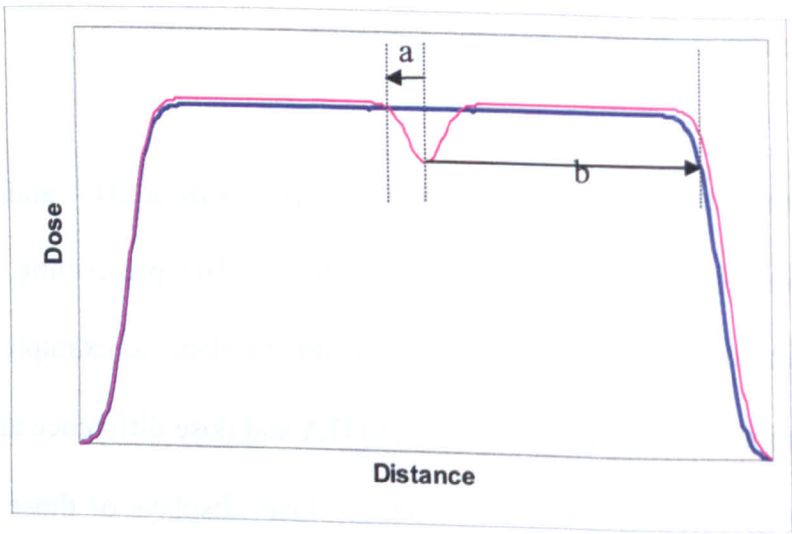



**Figure 3.2:** Sketch example of the asymmetry of the DTA evaluation of two dose profiles. The arrow *a* shows the distance from the central dose on the blue line to the nearest same-dose point on the pink line. The arrow *b* shows the distance from the central dose on the pink line to the nearest same-dose point on the blue line.

treatment, and thus the person analysing the data would have to attempt to juggle the two displays to see why each fails in certain areas.

Another issue with DTA is that the function is not symmetric. If the ideal distribution is compared against the non-ideal distribution, a different answer may be obtained than if the non-ideal is compared against the ideal. This is shown graphically in figure 3.2. The blue ideal distribution is a calculation of the dose profile across a single 6 MV X-ray beam as measured 5 cm deep in water. The shape of the beam penumbra is such that the same profile could be created by using two smaller fields abutting each other in the centre of the profile. Also shown on the graph is a non-ideal profile (pink), representing a possible measured profile. The large dip in the centre of the profile could be caused by badly aligned abutting fields. Analysis of the DTA for the central point may yield two answers: If the analysis takes the non-ideal dose and finds the nearest point in the ideal profile with the same dose, the distance is *b*; if the analysis takes the ideal dose and finds the nearest point in the non-ideal profile with the same dose, the distance is *a*. If the dose anomaly in the non-ideal profile is a spike rather than a dip, and the spike is higher than any point in the ideal profile, then the distance *a* becomes infinite – no dose match is found. The locations of calculated values of DTA also depend on which order the DTA calculation is performed. Referring to the steep gradient right-hand-side in figure 3.2, there is a DTA of 2 mm along the penumbra, but the calculated values of DTA could be recorded at the points along the blue line or along the pink line, and thus the recorded values of DTA would be 2 mm apart depending on the direction of calculation. There is clear scope for the DTA values to be placed at the wrong place in the answer matrix and therefore correspond to the wrong part of the original dose distribution.

**Composite index**

In reality, dose distributions are never simply shifted spatially or in dose, but a combination of each over different areas. In an attempt to combine both the dose difference and DTA measures into a single analysis method, Harms *et al* (1998) developed the Composite index. The method for calculating the Composite index is as follows: Binary matrices of data are calculated for DTA and dose difference, based on pre-selected tolerances (for example, 3% and 3 mm) such that a 1 in the binary matrix indicates a point outside of tolerance. At every point across the DTA binary matrix, the value is multiplied by the value for the corresponding location in the dose binary matrix, and the product is the Composite index value for that point. Thus a Composite index value of 1 for a point shows that the point falls outside both the DTA and dose difference tolerances, whereas a 0 shows that it passed at least one tolerance level. For Composite index values of 1, the magnitude of the dose difference at that point is then assigned. This method therefore takes the best elements of the DTA and dose difference binary matrices and will automatically select the DTA measure for steep dose gradients and the dose difference for shallow dose gradients. The Composite index is really only a qualitative measure of how well the ideal and non-ideal dose distributions compare to each other, because it shows only pass/fail for areas of the analysis and provides no numerical data. However, it was the first time that DTA and dose difference were combined in one index and begins to circumvent the issues apparent with each measure. The Composite index suffers from the asymmetry problems of the DTA method as DTA is one of the components of the calculation process.

**Gamma index**

In the same year, Low *et al* (1998) developed the first quantitative measure of dose distribution congruence that combines the DTA and dose difference methods into a single index. Called the Gamma index, the method is similar to the Composite index, but uses a vector combination of DTA and dose difference to provide a value that must pass a single tolerance. The numerical values thus returned from the Gamma index provide a more continuous measure of the dose comparison than the binary results of the Composite index.

The Gamma index at a point $\mathbf{r}_i$ in an ideal distribution is defined as (Low and Dempsey 2003):

$$\gamma(\mathbf{r}_i) = \min\{\Gamma(\mathbf{r}_n, \mathbf{r}_i)\} \text{ for all } \mathbf{r}_n$$

where

$$\Gamma(\mathbf{r}_n, \mathbf{r}_i) = \sqrt{\frac{r^2(\mathbf{r}_n, \mathbf{r}_i)}{\Delta d^2} + \frac{\delta^2(\mathbf{r}_n, \mathbf{r}_i)}{\Delta D^2}}$$

$$r(\mathbf{r}_n, \mathbf{r}_i) = |\mathbf{r}_n - \mathbf{r}_i|$$

$$\delta(\mathbf{r}_n, \mathbf{r}_i) = D_n(\mathbf{r}_n) - D_i(\mathbf{r}_i)$$

That is, the Gamma index, γ, at a point of interest $\mathbf{r}_i$ in the ideal dose distribution is the minimum of a set of test Gamma values, Γ, which are calculated for the point $\mathbf{r}_i$ and all points in the non-ideal distribution ($\mathbf{r}_n$). An individual Γ value is calculated for the points $\mathbf{r}_i$ and $\mathbf{r}_n$ by taking the square of the difference, $\delta$, of the dose $D$ at points $\mathbf{r}_n$ and $\mathbf{r}_i$ and dividing by the square of the dose difference tolerance $\Delta D$. This is then added to the square of the distance $r$ between $\mathbf{r}_n$ and $\mathbf{r}_i$ divided by the distance tolerance $\Delta d$ and the whole is then square rooted. The calculation is repeated for all of the points in the non-ideal distribution to produce a plane of Γ values centred on the point $\mathbf{r}_i$ in the ideal distribution, and the minimum value of Γ is selected as γ for the point $\mathbf{r}_i$.

By dividing the distance and dose difference by the respective tolerances within the square root, the values are normalised and the γ index is the addition in quadrature

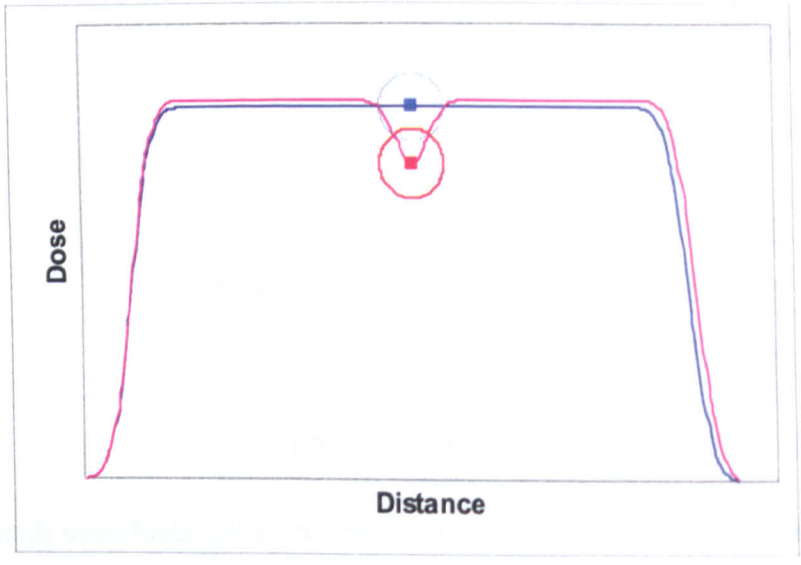



**Figure 3.3:** Sketch of the asymmetry of the Gamma index. If the Gamma index is evaluated at points along the ideal dose profile (blue line) then the non-ideal profile (pink line) will always pass through the Gamma acceptance circles. If the Gamma index is evaluated at points along the pink profile, there are places where there is no intersection of the blue profile and the red acceptance circle.

of dimensionless dose and distance measures. A Gamma index of less than 1 for a point $\mathbf{r}_i$ therefore corresponds to that point being within whatever dose and distance tolerances are set. Graphically speaking for a 1-dimensional dose profile, the Gamma evaluation method draws around the point $\mathbf{r}_i$ in the ideal distribution an acceptance ellipse whose axis radii are determined by the tolerances set. Dividing the dose and distance by the tolerances changes the scale of the axes such that the acceptance ellipse becomes a circle of unit radius. The Gamma index for the point $\mathbf{r}_i$ is then the distance from the centre of the circle to the closest data point in the non-ideal distribution. If any point in the non-ideal distribution falls within the circle, the Gamma value at that point is less than one and the point passes the test. In higher dimensions, the Gamma acceptance envelope becomes a sphere for 2-dimensional dose distributions or a 4-dimensional hypersphere for 3-dimensional dose distributions. In general, the acceptance envelope is of dimension N+1 for an N-dimensional data array because the space in which Gamma operates has N spatial and 1 dosimetric dimensions.

The process of minimum-selection means that the Gamma evaluation method automatically finds the best agreement between the two dose distributions from the DTA (where a test point $\mathbf{r}_n$ is found with zero dose difference), dose difference (at the central test point $\mathbf{r}_i=\mathbf{r}_n$ where the distance is 0) or a vector combination of both.

The Gamma index, like the Composite index and DTA, is not symmetric in terms of the analysis order of the two dose distributions being compared. Figure 3.3 demonstrates this for the 'spiked' dose profile from figure 3.2. If the depth or height of the spike is greater than its width and the width is less than twice the DTA tolerance of the Gamma index, then there is a potential for the Gamma index to ignore the spike. For example, a circle is shown in figure 3.3 around the central point of the blue ideal profile, representing the Gamma acceptance circle for evaluation along the ideal profile. Because the width of the pink underdose spike is small, the edges of the spike fall

within the blue acceptance circle. If the evaluation is carried out along the pink non-ideal profile then at the bottom of the spike there is no agreement of the blue dose distribution and the red acceptance circle, and the trouble spot is highlighted. Low *et al* defined the Gamma index as being evaluated at points along the ideal profile and in its original formulation the Gamma index may therefore miss critical features in a non-ideal dose distribution. This issue is similar to the asymmetry with the DTA measure on its own shown in figure 3.2.

The situation is reversed for situations in which a 'spiked' dose profile is the ideal dose and the non-ideal dose is a flat profile. In this situation, using the formulation of Gamma proposed by Low *et al* will highlight the missing spike whereas the reversed formula would ignore it. This issue will be discussed further alongside the other dose metrics later.

The Gamma index thus provides a single index of information which neatly combines the DTA and dose difference without the issues associated with high and low dose gradient areas. Care must be taken in analysing the results of the Gamma index to ensure that small areas of high dose difference are not ignored. However, the Gamma index has become one of the standard tools for analysis of dose in clinical environments (Chung *et al* 2005, Cernica *et al* 2005, Duthoy *et al* 2004) due to the ease with which it can be implemented and the ability to compare results between centres.

### Multi-pass Gamma

The Gamma index method operates on and utilises only discretely sampled data. That is, the Gamma index for a selected data point on an ideal distribution is the normalised distance to the nearest data point in the non-ideal distribution. As such, the Gamma index can be artificially higher than it may need to be for certain cases of dose distribution. Depuydt *et al* (2002) described these cases and proposed a refinement for

the Gamma index calculation method, based on a multi-pass approach. The first pass operates in exactly the same way as the conventional Gamma index. The subsequent passes operate only on points that have failed previous passes of the algorithm. The second and third passes effectively interpolate between the discrete sample points to find points along the lines that join them that may lie inside the acceptance ellipsoid. The method does not calculate the distance to these interpolated lines, only if they pass through the ellipsoid or not. Thus the refined Gamma method can only report a binary pass-fail output, although every point is accurately reported for its pass or fail, unlike the original Gamma evaluation.

**Chi index**

Bakai *et al* (2003) proposed an alternative refinement to the Gamma calculation algorithm, the Chi index. This not only retains the numerical value of the index, but also helps to solve one of the perceived problems of the Gamma index, namely the calculation time. In searching for a minimum, the original Gamma index method involves calculating many test values of $\Gamma$ that are discarded – for analysis of a 20×20 $cm^2$ area at 1 mm pixel spacing, there are 40,401 calculations of test Gamma values ($\Gamma$) for each Gamma index value ($\gamma$) that is retained. This is then repeated for each of the 40,401 locations in the array of the ideal distribution, therefore requiring over 1.6 billion calculations of test $\Gamma$ values. Each calculation requires many floating point operations, and thus an un-optimized Gamma index analysis can take a very long time to compute - Stock *et al* (2005) report that the calculation of Gamma index for a 140×170 $mm^2$ area requires 15 minutes on a typical gigahertz PC. Although in general the search space for the Gamma calculation could be reduced to a smaller area around the interest point, the possibility arises of incorrect calculation of the Gamma index if search reduction is implemented without thought.

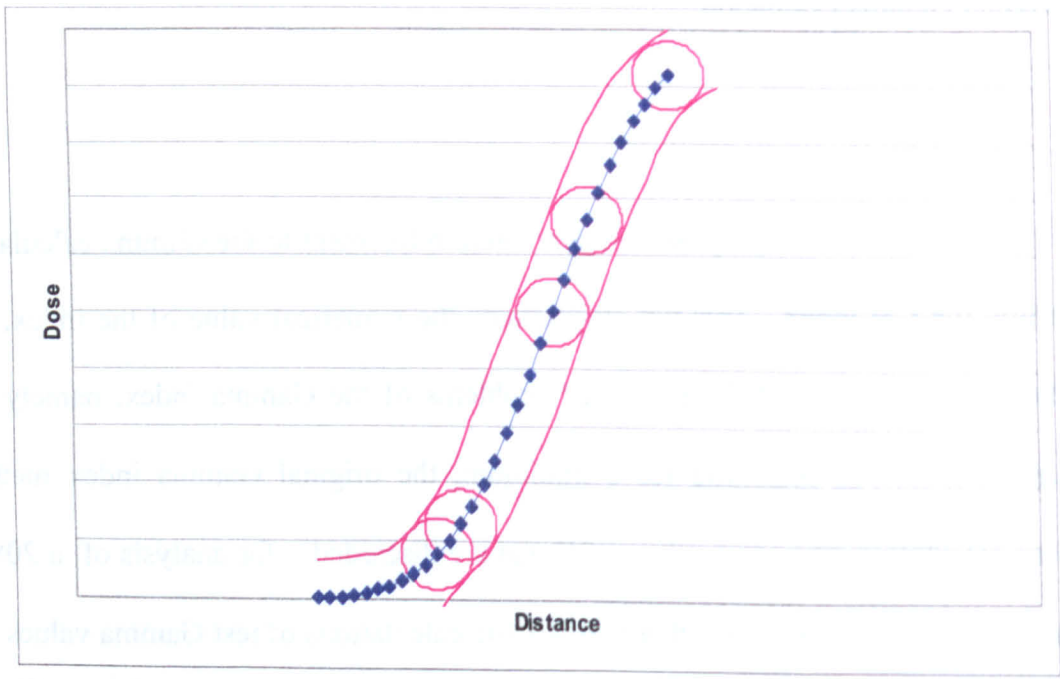



**Figure 3.4:** Graphical representation of the acceptance tube defined by Bakai *et al*. The tube is the envelope of the acceptance circles (both pink) created around each data point on the dose profile (blue) by the Gamma index

Bakai *et al* identified that where the Gamma index method draws ellipsoids around each discrete data point, the continuous analogy is to draw a tube around the dose profile (for one-dimensional analysis) inside which the non-ideal dose distribution must lie (see figure 3.4). The radius of the tube is the same as the radius of the acceptance circles used for Gamma, which is 1 in normalised dose-distance space. In higher dimensions, the tube analogy becomes harder to visualise, but the same principle applies of creating a continuous envelope rather than separate circles or spheres. The envelope for the Chi index is created by applying variable local dose tolerances at each point. Each dose tolerance is calculated as the vertical height of the tube based on the assumption that the dose profile has a constant gradient. The local dose tolerance is therefore proportional to $(\cos \alpha)^{-1}$ at every point where $\alpha$ is the gradient in % per mm. Once the local dose difference tolerance is computed for all points in the dose distributions, the Chi index is simply calculated as a fraction of the acceptable dose difference at every point and is thus very fast. Bakai *et al* report that this calculation method takes less than 1 second and is some 120 times faster than a 7 mm reduced-search Gamma index (25 seconds), 4000 times faster than a full-search Gamma index (15 minutes) and produces very similar, although not the same, results to the original Gamma evaluation. The small differences in the results of the Chi and Gamma indices are related to the method of calculation of the local dose differences based on the assumption of constant dose gradient. The Chi index does not address the asymmetry issues associated with any of the described dose evaluation metrics, as it is simply the continuous analogy of the discrete Gamma index.

**A new method for optimizing the speed of the Gamma calculation**

Many of the proposed refinements to the original Gamma index evaluation are based on the requirement for speed of calculation. As described above, the mathematics

of the Gamma index as given by Low *et al* require over 1.6 billion calculations steps for a 200×200 analysis grid, with each step requiring many floating-point operations. With a typical gigahertz PC, a calculation time of 10 to 15 minutes for this evaluation has driven the search for optimization and refinement.

All of the refinements proposed have made changes to the original Gamma index formulation. However, with suitable optimization of the computer programming code used, the Gamma evaluation can be calculated quickly, accurately, and in its original form. The Gamma index must be calculated for all points of the analysis area, so optimization must come in the form of reducing the search space for the minimum value of Gamma at each point. A simple method is to restrict the search area. While in general there is a possibility that we may alter the outcome of the search by restricting the search space, by setting the search area sufficiently large, we can ensure that we cannot find smaller values of Gamma outside of the search area.

A simple restriction is to reduce the search area to a radius of three times the DTA tolerance distance. With this approach, there is a small probability that a point outside of this area may yield a smaller value of Gamma than any inside it. However, due to the vector method of addition of the DTA and dose difference, the minimum value of Gamma found outside of this area will be 3, even if the dose difference is found to be 0. Clinical uses of the Gamma index method will require as much of the area as possible to have an index of less than 1, so it is safe to assume that the analysing physicist is not concerned with the actual value of the Gamma index if it is larger than 3, only that it is larger than 3 and therefore an unacceptable fail of the test. As an example, the smallest value of Gamma inside the reduced search space may be 3.5. There could be a point of dose agreement just outside the search space, which would therefore give a Gamma index of 3.1. The value returned by the reduced search would be 3.5, and this would be capped to 3 showing that the point being analysed fails the

Gamma index badly. If there is a value of Gamma inside the search area less than 3, then it will always be correctly returned as the minimum value. For a 1 mm data point spacing and a 4 mm DTA tolerance, this search-reduction method reduces the search area to a 12-pixel radius, or 450 data points, a reduction of a factor of nearly 100 over searching an entire plane of 40,000 points.

An additional method, similar to the above, is to restrict the search space to a radius based on the dose difference at the point for which the Gamma index is being calculated. The dose difference at the point being analysed can be converted to a distance using $r=(\delta/\Delta D)\times\Delta d$ where $\delta$ and $\Delta D$ are the dose difference and difference tolerance and $\Delta d$ is the DTA tolerance. In other words, we can restrict the search area to a radius that would have the same Gamma value for pure DTA as the central test point has for dose difference. Outside of this area all points will have a Gamma index higher than the central point, so can safely be ignored. In order to make sure that the search area is always big enough on a discretely sampled dose distribution, the radius should always be rounded up to the next whole data point spacing and one extra data point added.

Adding this method to the fixed search restriction above, we can select the smaller of the two for every point. Therefore for points with large dose differences, we restrict the search to three times the DTA tolerance and for small dose differences we restrict to the converted distance as described.

These improvements to the Gamma search algorithm do not change the way the algorithm is calculated, and the calculation problem remains an $O(n^4)$ calculation, where n is the number of data points across a plane or the reciprocal of the data point spacing. If we double the resolution of the data point arrays, there are 4 times as many points to calculate Gamma for, and 4 times as many points need to be checked for each. Improvements to the speed of the calculation must therefore be made in the minutiae of

programming coding. For example, the distance part of the square root in the Gamma calculation $\sqrt{\frac{r^2}{\Delta d^2}+\frac{\delta^2}{\Delta D^2}}$ need not be calculated for every test value of Gamma if the data is sampled on a uniform rectilinear grid. The distances between pixels will always be a simple function of their locations, and the distances from a central point to pixels nearby can be calculated once and stored in an array to be accessed later. Equally, the squaring of the tolerance values can be performed once instead of for every iteration of the calculation, and the square root only needs to be applied once since the square root of the minimum of a set of numbers is the same as the minimum of the square roots of the numbers, at least for the real positive numbers found in the Gamma evaluation.

These simple optimization routines have been implemented into a program and the calculation of the Gamma index can be performed for a 201×201 (1 mm×1 mm data spacing) grid of data on a gigahertz PC in around 1.8 seconds for a DTA tolerance of 4 mm and a dose tolerance of 4% of the maximum dose. The calculation methods of other authors were also implemented and the calculation times were found to be in agreement with those quoted by the authors. The implementation of the improvements detailed here is a significant improvement on the 25 second calculation time that Bakai *et al* reported for a reduced-search Gamma calculation, far better than the 15 minutes reported by Stock *et al* (2005) and also faster than the 3-4 seconds for their more complex optimization. The time of 1.8 seconds is more than Bakai *et al*'s tube method, which requires only 0.2 seconds. However, if a calculation takes less than 2-3 seconds, improvements in speed are almost unnoticeable to the user if the process is not performed repeatedly in quick succession. Thus the need for refinements to the Gamma index that trade increased speed for lessened accuracy or which modify the original Gamma index formulation is eliminated with a fairly simple set of optimisation routines which are easily implemented.

**Proposal of an accurate Gamma calculation method**

The Gamma index as originally proposed by Low *et al* suffers from the problem that it operates only on discretely sampled data and that there is therefore the significant possibility that areas of dose distributions that are in agreement may be reported as failing tolerance. Bakai *et al* and Depuydt *et al* both present methods to allow the Gamma test to operate on more continuous data and the interpolation of data between sampled points. Both methods presented raise their own issues – Bakai *et al*'s Chi index fails to accurately calculate the local dose tolerance as it is based on the assumption of constant dose gradient and Depuydt *et al* reduce the Gamma index to a binary pass/fail index. What follows here is a proposal for the development of a Gamma index calculation method that includes continuous/interpolated data without the issues of the Chi index or multi-pass Gamma method.

Bakai *et al* described calculating a local dose tolerance at each point of the ideal dose distribution, although it was based on constant local gradients. The concept of an acceptance envelope can be expanded if the envelope is created by performing a 'growing' or dilation algorithm to the dose distribution. The 3-dimensional analogy (which equates to a two-dimensional dose distribution because one dimension is taken up with dose) is the same as the way that 3-dimensional structures are grown to add margins in radiotherapy treatment planning (see chapters 1 and 2). The 2-dimensional dose distribution can be viewed as an undulating sheet in three dimensions. This can then be grown to create an envelope with a vertical margin specified by the dose tolerance and the horizontal by the distance tolerance. If the dose evaluation were to be performed in 3+1 dimensions, the envelope becomes a 4-dimensional object, but the basic approach holds. In terms of creating an envelope of dose tolerance, the growth technique is relatively easy to implement in a radiotherapy environment since the

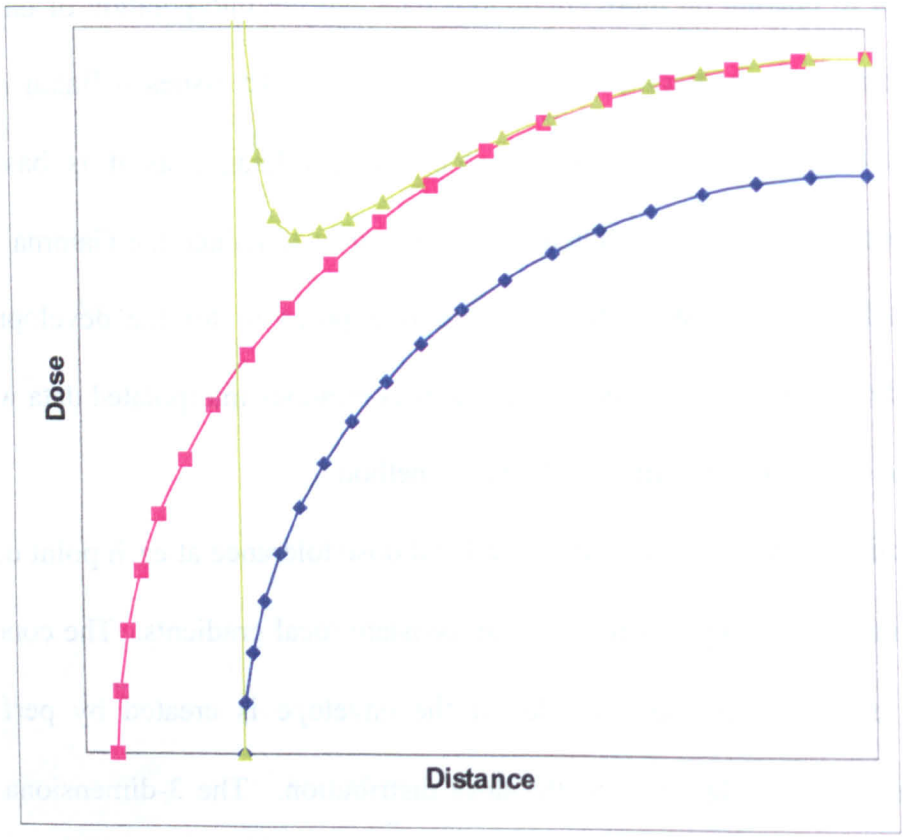



**Figure 3.5:** Schematic representation of the difference between the Chi index acceptance envelope and that of Gamma. The blue line represents a dose distribution, and the pink line shows the Gamma acceptance, a constant distance from the dose profile. The green line shows the Chi acceptance envelope, which is calculated based on local gradient and assumes a constant dose gradient. When the dose profile reaches infinite gradient (which is physically impossible), the difference between gamma and chi becomes infinite also.

algorithms already exist for treatment planning. The 3D growth technique does not make assumptions about local gradients and creates a true envelope where the 'thickness' of the envelope is always 1 in normalised dose-distance space.

The above describes how to create an accurate dose envelope for use with Bakai *et al*'s local dose tolerance method of Gamma evaluation. However, the local dose concept is still flawed. The local dose method returns the dose difference as a fraction of the vertical height of the dose envelope at a point of interest, and equates that to the Gamma index. The Gamma index, however, returns the shortest distance between the point of interest and the non-ideal dose profile. For areas where the gradient of the ideal or non-ideal dose profile are changing the two concepts do not equate, and figure 3.5 shows this graphically.

The concept of the Gamma index is to calculate the shortest distance from the ideal dose point to the non-ideal dose distribution, and that concept is presented here. Imagine a 2-dimensional non-ideal dose distribution as a 3-dimensional landscape, where vertical height represents dose and horizontal distances represent distance within the dose distribution's plane. A point in the ideal distribution for which we wish to calculate the Gamma index will, in general, sit either above or below the landscape. Once the vertical and horizontal axes have been rescaled to account for the dose and distance tolerances used for the Gamma index, the Gamma index is the shortest distance from that point to *any* point on the 'ground'. If the non-ideal dose distribution were sampled at an infinite resolution, then this shortest distance could be found by finding the distance from every point on the surface to the ideal dose point and taking the minimum. However, the non-ideal dose distribution will always be discretely sampled, usually on a rectilinear grid.

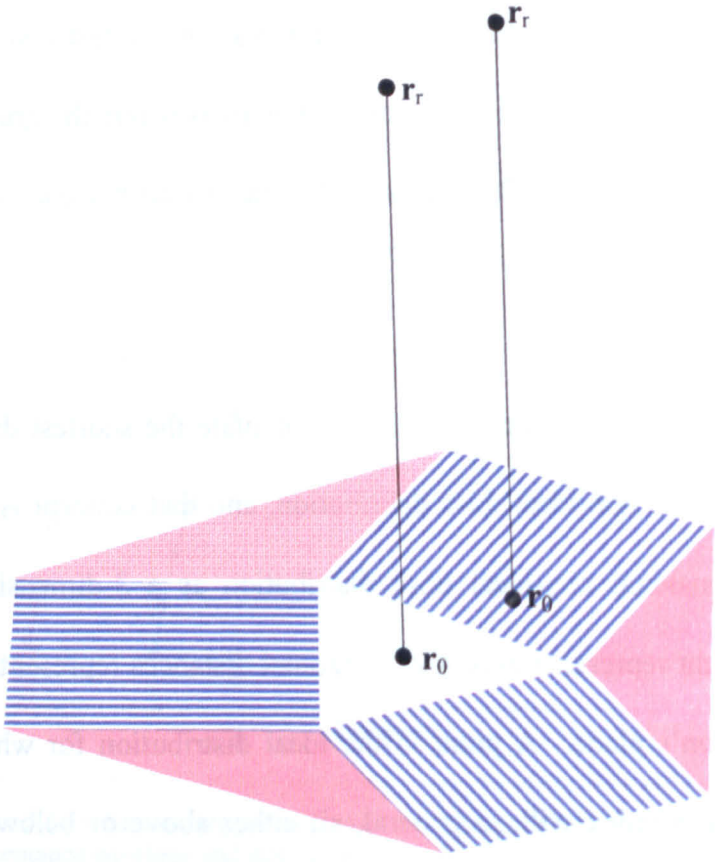



**Figure 3.6:** Representation of the shortest distance from a point $\mathbf{r}_r$ to the white triangle in three dimensions. The perpendicular from $\mathbf{r}_r$ to the plane of the triangle intersects the plane at point $\mathbf{r}_0$. This point may fall inside the triangle or in one of 6 regions around the triangle. In each region, the closest point of the triangle to $\mathbf{r}_0$ and $\mathbf{r}_r$ is either an edge (blue striped region) or a vertex (pink region).

The shortest distance can then be calculated thus:

1) The sampled dose distribution 'landscape' is divided into a series of triangles, usually by dividing each rectangle defined by the grid into two right-angled triangles by drawing a diagonal line or into four by calculating an interpolated central point.
2) The distance to each triangle is calculated and the minimum of all distances is taken:
    a. The distance to the plane of a triangle is calculated. If the perpendicular line from that plane to the point passes through the triangle, then the algorithm stops as this is the shortest distance to any part of the triangle. If the perpendicular falls outside the triangle, it passes through one of 6 regions (as shown in figure 3.6) and the closest point on the triangle is either an edge or a vertex, and this distance is calculated. The minimum of the plane, line or vertex distance is returned
    b. The distance to all triangles in the vicinity of the test point is calculated and the minimum recorded, which is returned as the Gamma index for that point.

This calculation method provides the shortest possible distance from the point being tested to the dose surface itself, not just the sampled points. The original formulation of Gamma only measured distances to points, and subsequent methods to interpolate between the sampled points have led to small inaccuracies (such as with the Chi method) or reduction of the index to pass/fail (as with the multi-pass method).

The calculation method proposed here will only calculate distances to planes between sampled data points, which is linear interpolation. For a curving surface, this

may be slightly inaccurate as the real dose surface may be concave or convex between the sampled points. This method does not allow for curved interpolation methods such as bicubic or spline during the calculation, as the distance function relies on distances to planes defined by 3 points. If there is a need for the accuracy of the method to be improved beyond that which is achieved with this method, a curved interpolation method could be used to resample the original data before the Gamma calculation is performed.

The Gamma calculation method proposed here is likely to be slow, as it requires several calculations per data point, of which a typical Gamma plot has 40,000. If the data is interpolated, then this number increases with the square of the data point spacing. However, implicit in the algorithm is searching only near the point of interest. When the dose planes are viewed as three dimensional surfaces, it becomes clear that no closer points can be found on the surface at a radius from the point greater than its height above or below the plane. Thus simple optimization (as suggested above for the normal Gamma calculation) could significantly increase the speed.

**Kappa index – a new dose evaluation metric**

To complement the dose evaluation metrics described above, a new metric is introduced in this thesis. This index will be called the Kappa index, $\kappa$, and is based on the Composite index of Harms *et al*, but retains the numerical nature of indices such as Gamma and Chi. The Kappa index shows at every point for a non-ideal dose distribution, the minimum of either the normalised dose difference magnitude or the normalised DTA for that point. This differs from the Composite index not only in the presentation of a numerical index for every data point in the output matrix, but also in the methodology of the calculation. The Composite index does not need to calculate the actual distance to agreement at any point, and this is why it cannot present numerical

values based on it. The calculation for the Composite index searches over an area for a dose agreement point, the radius of that area being the distance to agreement tolerance. If dose agreement is not found in that search area, then the DTA tolerance fails for that point. The Kappa index *requires* the calculation of the distance to agreement matrix for every data point in the ideal distribution so that it can then be compared against the dose difference at the point of interest.

This method of either-or dose evaluation is difficult to implement due to the difficulty of calculating the DTA matrix for a pair of dose distributions. The method described by Childress and Rosen (2003) is easy to understand and simple to implement with the right tools. Their method involves pre-computing a set of isodose contours for both dose distributions at the beginning of the calculation. For each calculated point on an isodose in the ideal distribution the shortest distance to the same level isodose in the non-ideal distribution is found. The resulting data is then interpolated to a rectilinear grid. Childress and Rosen report good results when using 3 cGy isodose intervals. However, there is the potential for an isodose in the ideal and non-ideal dose distributions to be quite far apart, so the approach for the Kappa index is slightly different. A set of isodose contours is calculated at 0.1% intervals for the ideal distribution, where 100% is chosen as the dose to the target volume. Each point in the non-ideal distribution is then analysed. The dose at each point is rounded to the nearest 0.1% and the distance to the corresponding isodose line from the ideal distribution is found. The MATLAB programming environment provides ideal tools for this process, as the built-in contour function returns contour vertices with sub-pixel accuracy and is also very fast. The closest point of approach of a point and a contour is found by calculating the distance from the point to each line segment of the contour and finding the minimum. This is also slightly different to the method chosen by Childress and Rosen, who calculate distance to vertices of the contour only, not to line segments,

although for finely-spaced data the differences are small. Typically each contour has around 100 vertices and thus the overall calculation of DTA requires approximately 100×200×200 = 4 million steps to evaluate over a 200×200 plane. This is some 400 times faster than an un-optimised Gamma evaluation and can be performed in around 4 seconds on a gigahertz PC running MATLAB. The pre-calculation of contours is a limiting factor to the speed of the DTA calculation process. If the contours are spaced too closely in dose value the pre-calculation will take a long time, but spacing them too far apart will result in a DTA plane that is not smooth. It was found that 1000 contours at 0.1% intervals provided a good resolution and could be calculated in around 10 seconds. Areas where the 0.1% contours are a long way apart have got very low dose gradients and the DTA is therefore less applicable in these areas anyway.

The Kappa index, reliant on the DTA, is also an asymmetric function, dependent on which dose distribution has the contours pre-calculated. Although Kappa can be calculated in either 'direction', the definition of the index is opposite to that of the Gamma index and its derivatives. In those methods, the index is evaluated at points on the ideal profile, looking for points on the non-ideal profile that are nearby. As discussed above, this leads to the possibility of ignoring spikes, either low or high, of dose. The Kappa index is evaluated at points along the non-ideal profile such that spikes of dose are not ignored. The asymmetry of the index means, however, that if a spike or trough of dose is *desired* and is not found in the measured dose distribution, then the possibility exists that the area will pass the Kappa index. However, as the next section shows, the Kappa index is less tolerant of errors in areas of changing dose and changing dose gradients.

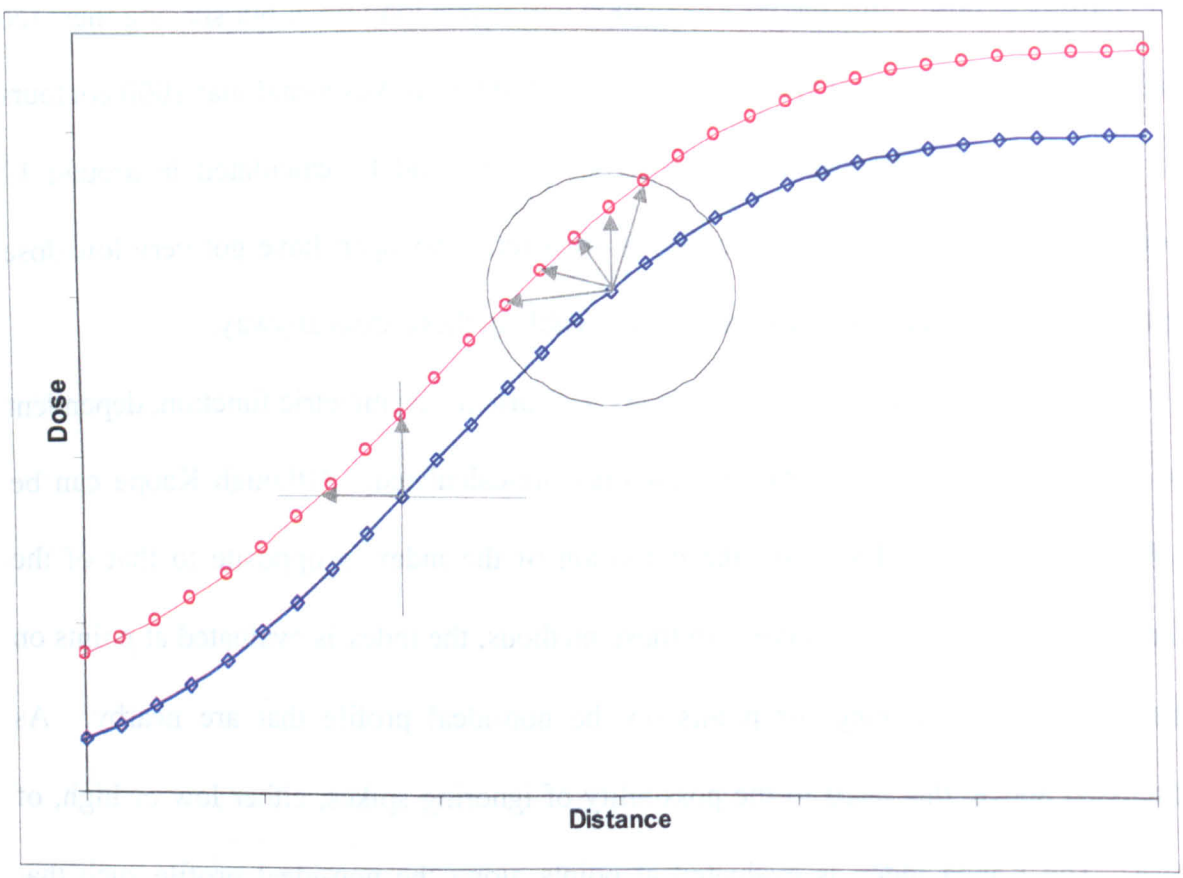



**Figure 3.7:** Graphical representation of the acceptance methods of the Gamma, Chi, Kappa and Composite evaluation methods. Kappa and the Composite index require that the non-ideal dose profile (red) passes through a cross centred on the ideal profile (blue), whereas Gamma and Chi allow the non-ideal profile to pass through a circle with a diameter equal to the size of the cross. The arrows represent evaluated distances in dose-distance space. Kappa returns the length of the shortest arrow on the cross, while Gamma returns the shortest arrow in the circle.

**Quantitative analysis of dose metrics**

The above described dose analysis tools – Gamma, Composite, Chi, multi-pass Gamma and Kappa – attempt to combine DTA and dose difference into a single index, albeit in slightly different ways. The different methods of combination of dose difference and DTA mean that each is slightly more or less tolerant of differing dose distributions than the next. Figure 3.7 shows graphically the acceptance methods of each index. The Composite index and the Kappa index methods draw an acceptance cross in dose-distance space. Each method demands that the non-ideal dose distribution passes through at least one spine if the point is to pass the test. The Composite index reports a pass-fail, the Kappa index reports the closest approach along either spine of the non-ideal dose profile. The Gamma index method draws an acceptance circle around each data point in the ideal distribution and that point achieves a Gamma index of less than 1 if any point from the non-ideal dose distribution falls within the circle. The Gamma index reports the closest approach of the sampled data points to the centre of the circle. The multi-pass Gamma method reports a pass or fail if any part of the non-ideal dose distribution passes through the circle or not, respectively. The Chi index method defines local dose tolerances either side of the ideal profile and reports the fraction of the dose difference at the point.

All of the methods combine DTA and dose difference in such a way that they reduce to being purely a measure of one or the other for horizontal and vertical dose profiles. For a vertical dose profile (i.e. infinite dose gradient), any small lateral shift of one dose distribution will produce a very large dose difference, so the DTA will always be reported as the 'winner' for each index. Equally, in a very low or zero dose gradient area, a small dose change will produce a large DTA and the dose difference will be returned. All of the indices thus produce the same results in these areas. With dose gradients that are neither zero nor infinite, the indices differ in their reported results, and

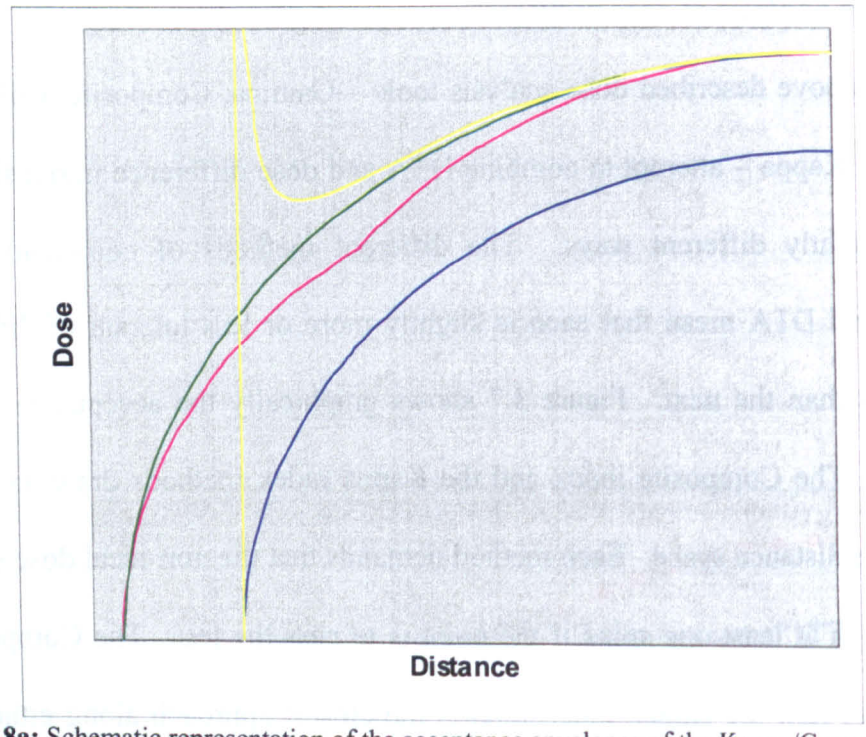



**Figure 3.8a:** Schematic representation of the acceptance envelopes of the Kappa/Composite index (pink line), Gamma index (green line) and Chi index (yellow line). The blue line represents a dose profile in 1 dimension. For low gradient areas the indices are the same, but diverge in moderate gradients.

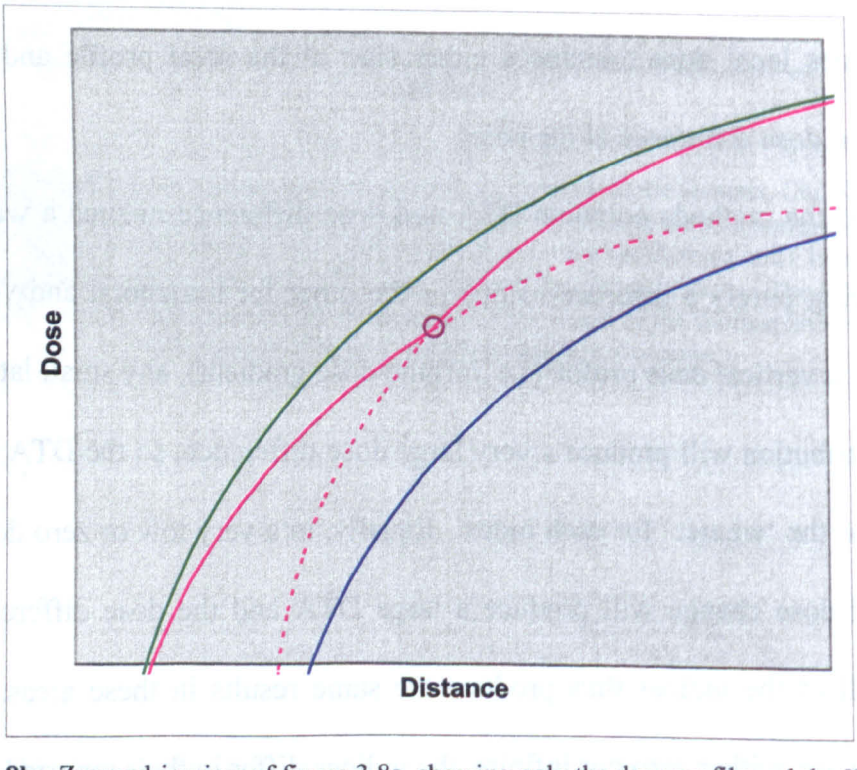



**Figure 3.8b:** Zoomed-in view of figure 3.8a, showing only the dose profile and the Kappa and Gamma envelopes. Also shown are the Kappa 'construction lines' of the original dose distribution shifted up or to the left. Note the kink in the solid pink line at the point indicated

this is especially so for non-uniform dose gradients. Figure 3.8a shows an example dose profile for which the indices differ in their results. Figure 3.8b shows a zoomed-in dose profile with just the Gamma and Kappa methods shown. From these diagrams, it can be seen that the Composite and Kappa index methods are the least tolerant of differing dose distributions, the Gamma method is more tolerant and the Chi is most tolerant, allowing dose points to pass that would fail the other index methods.

Figure 3.9 (over page) shows the case for a dose profile with a gradient of 1%/mm. This '45°' slope shows the greatest difference between the Gamma-type indices and the Kappa-style ones. The Gamma index allows the non-ideal dose to lie anywhere within a perpendicular distance to the ideal dose that is defined by the index's tolerances. So for a tolerance of 3% and 3 mm, the dose distributions can be up to 1 distance unit apart at 45° to each other in the rescaled axes of dose-distance space. The Kappa index and Composite index demand that the non-ideal dose lies within 1 distance unit either up/down or left/right of the ideal dose. For a 1%/mm dose gradient, this equates to the closest approach of the two distributions being $1/\sqrt{2}$ units apart, not 1 unit. Alternatively, the 1 unit distance that the Gamma index allows in the 45° direction allows a $\sqrt{2}$ distance in the dose-only or distance-only directions. What this means is that if an entire dose distribution is shifted by 3% *or* 3 mm (but not a mixture of both such as 1.7% and 2 mm), then the Gamma index will report values less than 1 when the dose gradient is non-zero, and the effect will be biggest for dose gradients closest to 1%/mm. If the dose tolerances and distance tolerances are not the same numerically then the effect will be largest for a different gradient, for example 1.5%/mm for a tolerance of 3% or 2 mm. If a dose distribution has a constant gradient across it, then the effects of the Gamma over-tolerance could be clearly seen. For a constant gradient of 45°, the entire dose distribution could be shifted up by up to $\sqrt{2}$ times the actual tolerance and the Gamma method would report an index of less than 1. For a 3%/3mm

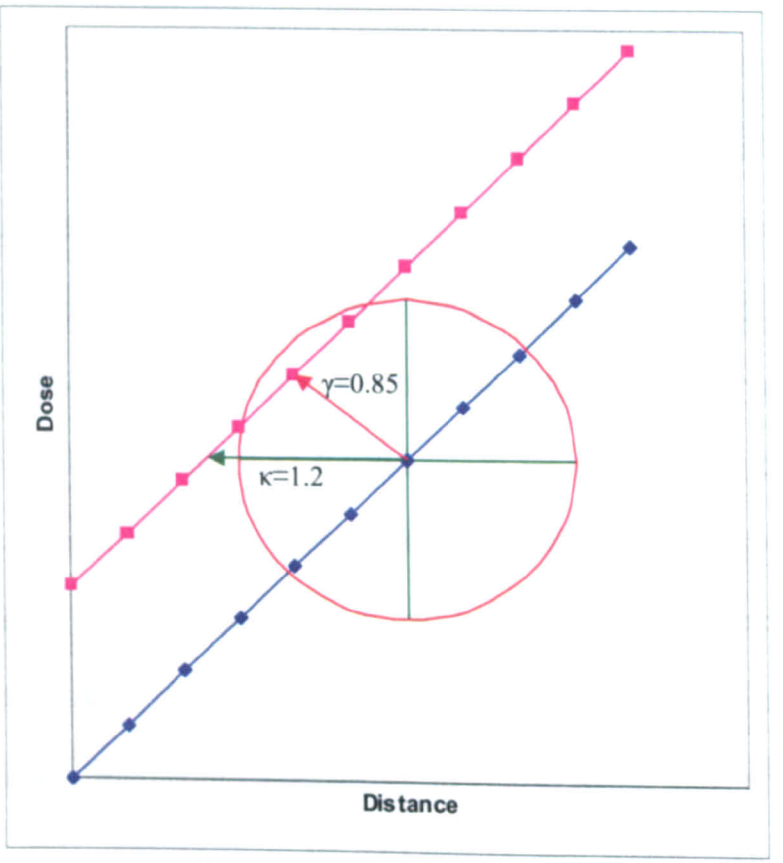



**Figure 3.9:** Schematic diagram of a 1% per millimetre dose gradient sampled every millimetre. The blue and pink lines represent two dose profiles separated by 3.6 mm (or 3.6%). The red circle shows the Gamma acceptance zone for 3%/3 mm, while the green cross shows the Kappa acceptance for the same tolerance. The pink dose distribution shows a minimum Gamma index of 0.85 whereas the Kappa index is 1.2 (shown by the red and green arrows respectively).

tolerance, this would allow dose or distance shifts of 4.2 mm. For a 5%/5mm tolerance, movements up to 7.1 mm or 7.1% would be tolerated by the Gamma index method. The Kappa and Composite methods are not over-tolerant in these areas, only allowing a dose change *or* a distance movement, and not a combination of both that equates to an increase in both.

For a typical dose distribution created within a patient during radiotherapy, and especially in IMRT, there are unlikely to be large areas of constant dose gradient that the Gamma method would be over-tolerant on. However, the dose gradient near the 95% isodose is typically around 1%/mm, and thus the Gamma method could be over-tolerant in this area. Treatment planning centres around the 95% isodose, ensuring that the whole PTV gets coverage by this dose to be certain that the CTV is treated to at least 95% of the prescription dose after setup errors and organ motions have been accounted for. The over-tolerance of the Gamma method to dose and distance deviations in this area could mean that coverage is not guaranteed. For carrying out treatment verification, the Kappa index would provide a better indicator of dose distribution conformance as it will show a direct distance or dose difference. Not only will the Kappa index highlight potential problem areas more readily than will the Gamma index, but the numeric values of the Kappa index are inherently easier to understand, as they are not a combination of dose and distance, but a simple presentation of only one. Assuming that the dose and distance tolerances have been set such that the relative differences in dose and distance are equally favourable, then the Kappa index provides a much simpler way of analysing dose discrepancies.

As a by-product of the Kappa index calculation, a distance to agreement array is generated, which can sometimes be useful in showing that an area of high Kappa index is in an area of low dose gradient. Although this will show that the dose difference is out of tolerance, a check on the dose distribution may show that the particular area is in

a low dose area and may enable the clinician to decide that the treatment is acceptable since the dose difference is in normal tissue and is a change from 10% dose to 20%. Although a dose difference plot is easy to generate, a distance-to-agreement plot is not, and would have to be separately generated for the Gamma index, but is included as part of the algorithm for the Kappa index.

**Conclusions**

The Gamma index has been useful in introducing quantitative dose analysis and in familiarising the radiotherapy community to the concept of single indices that can combine distance and dose differences. The Gamma index has brought great improvements to dose and IMRT plan verification. However, the concept behind the Gamma index means that it can be excessively tolerant of dose differences and spatial movements in dose distributions, and this is especially true in areas near the 95% isodose line. The original Gamma algorithm also operates only on discrete data sets, and attempts at interpolation have led to approximations and errors in calculation. The Kappa index, as proposed here, removes the over-tolerance and presents a simple dose or distance index that is easy to understand and robust in what it presents. An algorithm for the calculation of interpolated distance to agreement has also been presented, which allows for accurate Kappa indices to be calculated. The Kappa index is less tolerant of dose differences and may lead to more apparent failure of plan verification. However, it is better that a dose distribution that is close to tolerance be flagged up for evaluation by a human than for an algorithm to blindly allow it, and the Kappa index can therefore only improve the accuracy of those plans that are passed, and by implication therefore patient outcomes.

Although a different dose evaluation metric has been presented here, it would be optimistic to expect that radiotherapy centres will immediately switch to using it. The

Gamma index is well-used around the world and is therefore a standard with which radiotherapy centres can compare results. As such, this thesis presents some improvements to the algorithm that address some of the issues that surround it. A method for the simple optimization of calculation speed has been suggested, along with a proposed method, although not implemented, of calculating an accurate, interpolated Gamma index that is not reduced to a pass/fail measure as some interpolation algorithms have done.

By improving the accuracy of calculation of measures in use such as the Gamma index, improvements in the analysis of dose distributions can be achieved. This can then help in the analysis of the causes of dose errors and the eventual reduction of errors. Reducing dose delivery errors in patient treatments can only lead to improvements in the planned outcome of their treatments, although by how much is an answer that remains unanswered. Work presented in chapters 4 and 6 aims to move towards those answers. Introduction of simpler more understandable dose evaluation metrics such as Kappa may also help to introduce physics-based analysis of dose distributions to clinicians and treatment planners, which may help to change prescribing and planning treatments to reduce errors and improve patient care.

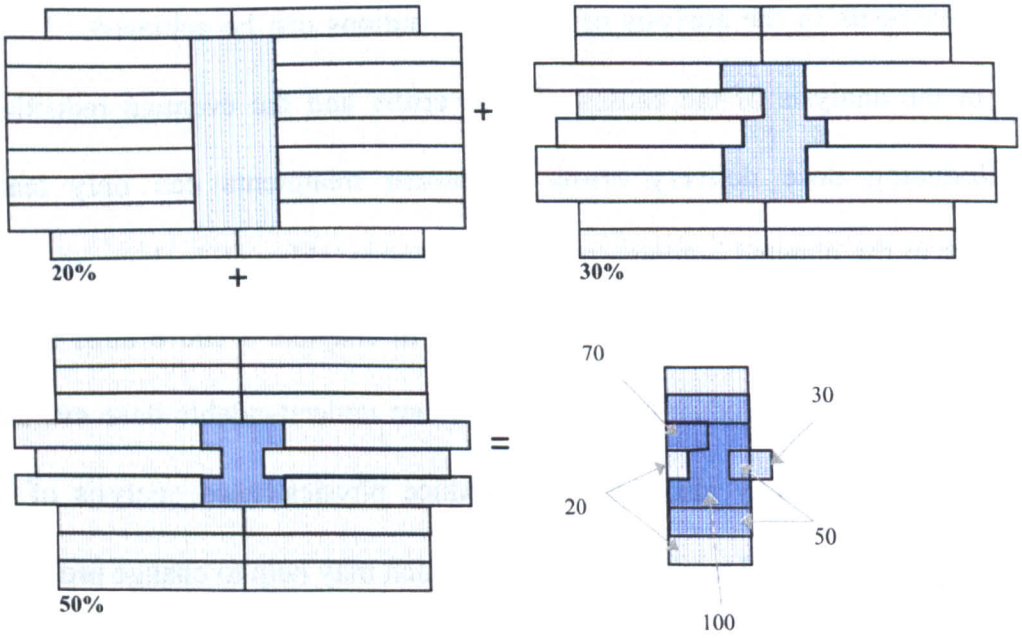



**Figure 4.1:** Construction of a simple fluence map by the summation of multiple MLC beamlet apertures. A large number of different intensity levels can be created with a small number of beamlets

# Chapter 4 – Systematic linear accelerator errors and their effects on IMRT dose distributions

## Introduction

IMRT using linacs is nearly always delivered using multi-leaf collimators to define beam aperture shapes. The multiple-static-field technique is commercially available from Siemens, Varian and Elekta. IMRT delivered using the MSF technique from a Siemens linac is used at Addenbrooke's.

MSF IMRT works by irradiating several beamlets (or segments) of different MLC-defined shapes to build up a modulated fluence map for each field. In this way, the fluence of radiation across a delivered field will be deliberately uneven (see figure 4.1). As briefly described in chapter 1, the method of operation for MSF beams is as follows: With the radiation beam turned off, the linac gantry is moved to the correct angle for the first beam and the MLC leaves are moved to define the first segment's aperture shape. When the MLC leaves are in position, the radiation beam is switched on to give the desired dose for that segment. When the segment's radiation exposure is completed, the MLC leaves move to the next shape and irradiation commences again. When all segments for a particular composite beam are completed, the gantry moves to the next position and the MLC leaves are moved to the first segment of the next beam.

In this way, any number of beamlets can be used to build up one field with an intensity modulated profile. Typical treatments use between 5 and 20 segments per beam. As a starting point, IMRT treatment planners may choose an odd number of beams (5, 7 or 9) spaced equally through the 360° rotation of the gantry to avoid directly opposed beams. The beam angles may then be adjusted to avoid direct irradiation of radiosensitive structures. Opposed beams hamper the inverse planning process used for most IMRT planning as the optimization routines used by the computer planning system fail to decide properly how to distribute radiation between the

beamlets. Frequently, the planning system will produce beams with very highly modulated fluences, the troughs in one beam being filled in by the peaks in the opposing beam. This degree of beam modulation leads to a less desirable distribution of dose throughout the healthy tissue areas of the patient and is more prone to delivery errors caused by small spatial movements of the patient or machine. Conventional 3D-CRT planning uses different collimator angles to help fit MLC shapes to target volumes and organs at risk and to align any wedges used. IMRT treatment planners may choose to use the same collimator angles for all beams because the inverse planning system can use fluence modulation to shape dose distributions to the target volume. This can lead to the summation of 'tongue-and-groove errors' (Que *et al* 2004). Small changes in collimator angles between beams can reduce the summation of such errors, although the errors tend to be small at depth in patients or phantoms due to scattering of radiation. There is often therefore no major gain to be had by treatment planners changing collimator angles from the default for inverse-planned IMRT.

Three-dimensional conformal radiotherapy uses a small number of large-aperture beams (typically fewer than 6) to deliver dose to the target tumour volume, while IMRT treatments may use as many as 180 small-aperture beamlets to construct up to 9 fluence maps defining the radiation given to the patient. Thus IMRT treatments use a different range of aperture sizes and shapes than conformal radiotherapy, and are subject to different errors and constraints.

Like all mechanical equipment, linacs are not perfect machines, and linac manufacturers provide specifications of the positional and dynamic accuracy of their equipment. Varian, for example, quote MLC positioning accuracy of better than 0.2 mm projected to the isocentre, which is supported by independent studies (Chui *et al* 1996, LoSasso *et al* 1998). Siemens quote MLC leaf positioning accuracy of ±2 mm for

their Primus range of linacs and ±1 mm for their Oncor linacs. For conformal radiotherapy treatments using fields between 5 cm and 20 cm, 1 mm MLC leaf positioning accuracy equates to less than 2% of the beam aperture area, and is usually much less. However, for IMRT treatments using small fields that may be as small as 2 cm wide, a 1 mm error equates to 5% of the beamlet area, and up to 10% if opposing leaves are in error in the same (in/out) direction. Some IMRT treatments may even use beamlets with opposing MLC leaves as little 5 mm or 1 cm apart, in which case ±1 mm positional accuracy on the MLC leaves becomes a major factor in the accuracy of IMRT delivery. While it has been known for some time that accurate multileaf collimators are required for dynamic IMRT (Budgell *et al* 2000), step and shoot IMRT is considered to be a simpler technique and the effects of poor leaf accuracy on IMRT delivered statically are not as well known. Recent work (Zeidan *et al* 2004, Woo and Nico 2005, Stell *et al* 2004) has investigated MLC accuracy in step and shoot, but has mainly focussed on the errors within particular plans rather than the underlying causes.

MLC leaf positioning and verification is performed in a number of different ways, with each major manufacturer choosing a different approach. Varian MLC systems use a potentiometer attached directly to the MLC leaves to report feedback on their positions. In this way, the leaf-driving motors move the leaves until the potentiometer reports that the leaf is in the correct location. This direct connection between the MLC position and the control mechanism provides little scope for errors to be introduced other than the potentiometer itself. Varian therefore also utilise a self-calibrating system to reset the potentiometer calibration to account for any changes in its response. During re-calibration, leaves are driven slowly forwards towards a laser beam until it is broken by the leaf-end. The exact location of the leaf is then known at

that point and the potentiometer reading is recorded. In this way recalibration of the leaves is easily and quickly done, and can be performed every day.

Elekta MLC systems use a video-based system to record and feed back the MLC leaf positions (Jordan and Williams 1994). A video camera and mirror mounted in the head of the linac above the MLC banks continually watch markers located on the tops of the MLC leaves. The leaves are moved by the linac until they can be seen to be in the correct locations by the camera and linac control systems. Although there are possible problems with alignment of the camera system and optical recognition of the leaves, the system is easy to maintain and in general achieves good positional accuracy, quoted as better than 0.3 mm (Jordan and Williams 1994, Hounsell and Jordan 1997).

Siemens have employed a stepper motor and motor encoder setup for their leaf-positioning verification – the encoder records one click per step of the motor and the linac moves leaves into position by driving the stepper motor until the desired number of clicks has been recorded. The leaves are calibrated at certain locations along their travel range and the encoder values at these points are recorded into the linac. An interpolation between these values then enables the linac to position the leaves anywhere along their range. Thus the feedback from the leaves to the linac's control mechanisms are via the motor encoder and are not directly linked to the MLC – if the motor reports the MLC leaf as having been moved to the correct location then the linac records it as such. This mechanism leads to two problems with the leaf positioning accuracy. Firstly, if the encoder receives spurious clicks or misses any, the positional accuracy of the leaves will shift as the motor/encoder pair now report the MLC leaves to be somewhere other than their real locations. The second issue is that there is no direct recording of the leaf position, only that which the encoder reports. Thus any mechanical deficiencies in the MLC manufacture and design that prevent precise positioning will not be picked up by the linac control systems.

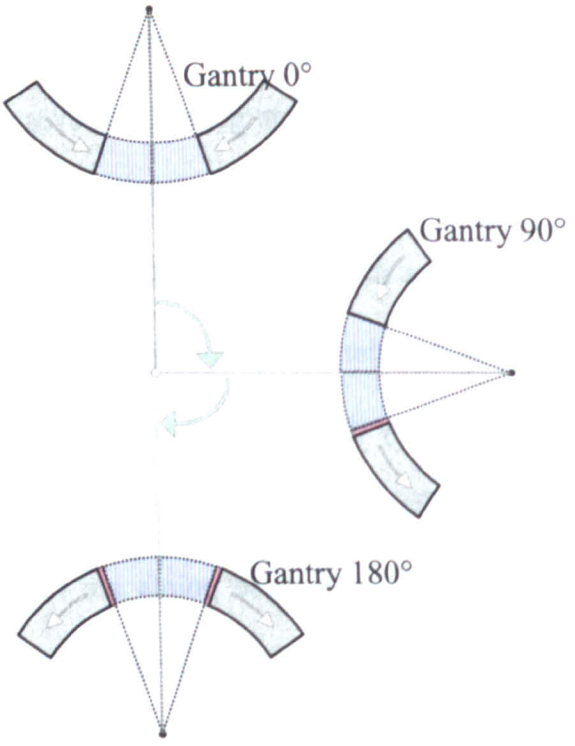



**Figure 4.2:** The Siemens MLC hysteresis problem. The MLC leaves are calibrated at gantry 0° with gravity (white arrows) pulling the leaves into the field aperture (blue area). When the gantry is rotated to 90° (green arrows), the top bank of leaves experiences the a gravitational force in the same direction and do not move. The lower bank of leaves experiences a force out of the field aperture and move, exposing extra field area (red area). At gantry 180°, both banks of leaves experience a force out of the field and extra field area is created on both sides of the field aperture.

The curved-track motion of the double-focused Siemens MLC leaves is a more complicated design than the linear-drive systems used by Varian and Elekta, and therefore mechanical tolerances are not as tight as they can be with the linear-drive systems. As such, there are small amounts of backlash in the Siemens gearing system, allowing the MLC leaves to move slightly without being driven, and also resulting in small differences in leaf position depending on the drive direction. These small changes may lead to systematic errors in MLC leaf positioning.

The MLC leaves may be calibrated by using the light field of the linac. The light field is a visual representation of the beam aperture, and is created by a mirror and light source in the head of the linac. The mirror and light are set up so that the virtual source of the light is nearly coincident with the radiation source, the difference between the two usually being known. The projection of the light field is therefore representative of the radiation given during beam-on. The MLC leaves can be moved using a computerised interface until the light field projection at isocentre shows them well aligned to a specific field size which is then recorded by the linac. The leaves are always calibrated with the gantry at the 0° upright position for ease of access to the light-field projection. The pendulum-like travel of the Siemens leaves means that at gantry 0° they will be exerting force on the control systems in a direction that tries to move the leaves into the field, and any backlash or hysteresis in the leaf movement will be taken up in this direction (see figure 4.2). This is the 'swung-in' position of the leaves. Thus if a pair of leaves are calibrated to give a 10 cm isocentric field width at gantry 0, the movement range of the leaf within its backlash range will be from 10 cm to $(10 + 2x)$ cm where $x$ is the amount of backlash on each MLC bank. When treatments are carried out at gantry 0° the field sizes will be correct, but systematic changes in field size will be observed when the gantry is rotated to other angles. As the

gantry rotates clockwise, for example, the X2 leaf bank (if the collimator is at 0°) will begin to experience less gravitational force on its leaves, until a critical point when the gantry is at approximately 10 degrees. At this point, gravity will be pulling on the leaves in the direction away from the beam centre and the backlash and hysteresis in the leaf motion will therefore be taken up in this direction. The forces on the X1 bank will still be into the field in the same way as at gantry 0°, so the leaves will be in the same location as calibrated. At this point, therefore, the field will be slightly larger than intended, with the increase in field size always on the lower bank of leaves, giving a field size of $(10 + x)$ cm. With the gantry at or near 180°, both leaf banks will be experiencing gravitational forces trying to swing the leaves out, and the field size will increase on both sides, leading to a $(10 + 2x)$ cm field size.

IMRT typically uses many low-dose beams as opposed to conformal radiotherapy's use of fewer, higher-dose beams. Thus IMRT places different demands not only on the MLCs but also on the performance and behaviour of a linac's radiation output. Radiation dose at a point in a patient is dependent on the radiation output of a machine, the source-to-surface distance of the patient, the electron density of material the radiation passes through and the depth of the point within the patient. Thus it is not possible to set a linac to deliver dose, and linacs are instead set to deliver a specified number of monitor units (MUs). Monitor units are a measure of how much radiation is measured by the internal dosimetry systems of the linac, and are thus designed to be proportional to the amount of radiation given out by the machine. Radiotherapy centres calibrate their linacs to deliver a specific amount of radiation per MU to a specific depth in water for a specific field size and source-to-surface-distance (SSD). For example, Addenbrooke's calibrates its linacs to deliver, at 100 cm SSD and for a 10 cm by 10 cm field, 1 Gy / 100 MU to a depth of 1.6 cm for 6 MV photons and 3 cm deep for 15 MV

photons (these depths correspond to those of maximum dose for each energy). The linac is thus calibrated to give 1 cGy / MU at the specified depth, SSD and field size. When set to run for more than about 10 MU, a linac will deliver reliably 1 cGy / MU. Below this value, variances in the internal measurement properties of the linac and 'ramping up' of electronics during the start of the beam mean that the cGy / MU ratio is close to, but not always unity. If these variances are stable over time then treatments that use beams with small numbers of MUs may be systematically affected as the treatment beams will fall at the same point in the non-linear behaviour for every treatment fraction. Additionally, because of the way that the intensity levels in some IMRT segmentations work, all segments for a particular beam may sample the same area of the linac's behaviour and thus all segments may be affected in the same way.

A linac's dose output also varies slightly from day to day with changes in humidity, temperature, atmospheric pressure and as components in the machine age and become less efficient in their measurement or operational properties. A program of regular quality assurance ensures that these variations remain within a small range of allowed values. Over the course of a patient's treatment, these variations will be random errors of the order of less than 1.5% (as measured at Addenbrooke's) and are therefore ignored in their effects on patient treatments, being absorbed into the various tolerances given on patient treatments.

Treatment planning systems and treatment planners do not account for imperfect machine parameters. That is, they rely on the machine having adequate device positioning accuracy and dose delivery characteristics. Thus the dose distributions calculated by treatment planning systems may not accurately reflect the dose distribution that the radiation will create within the patient, and this is especially true and more important for complex IMRT treatments. Although the overall dose

distribution may not appear significantly different, there may be discrepancies over small volumes that are clinically relevant, especially in the high dose gradient environment of closely-conforming isodose lines that IMRT treatments aim to create within patients receiving radiotherapy.

It is assumed that many of the issues that affect linac delivery characteristics are random errors: MLC leaf positioning errors may be one way or the other, dose linearity with MUs may be above or below 1 depending on the number of monitor units and dose output may vary slightly from day to day. Random errors should give a mean of zero error over the course of a fractionated patient treatment, but are difficult to quantitatively account for and quantify in their effects on a patient's treatment in terms of the standard deviation and spread of the errors (Thomas and Brooke 2002). Thus most deviations from perfect parameters that exist on linacs are kept within tolerances and their effects on the total radiation treatment of a patient are assumed to be negligible. However, as discussed above, if any of these effects become systematic errors then there is clear scope for treatment errors to occur over the whole course of a patient's treatment.

Because treatment planning systems do not allow for imperfect machine performance, it is very difficult to analyze the problems associated with either random or systematic errors in their behaviour. Previous work has shown how to combine random and systematic errors into treatment planning margins (BIR 2003), although this focuses mainly on the addition of distance errors associated with the many stages of treatment planning that affect the whole plan rather than the dosimetric differences created by small changes to parts of a treatment plan. Existing commercial treatment planning systems all assume constant linac parameters, exactly reproducing data measured or specified at the time of commissioning of the machines. Although

deviations in machine performance are expected, it is assumed that a thorough program of quality assurance checks on the linacs will keep them within small tolerances and that the deviations will be random rather than systematic. Treatment planning systems also do not allow linac data to behave differently as a function of parameter value, such as field sizes changing with gantry angle or dose output varying with numbers of monitor units. It is possible to produce treatment plans and modify some of their properties externally to the planning system and then re-import back into the system for recalculation. However, many treatment planning systems will not allow small modifications to plan parameters, for example only allowing integer numbers of monitor units or millimetre steps in MLC positions. Modifying plans to simulate 2% MU errors on beamlets of 10 MU or to move MLC leaves by sub-millimetre distances would therefore be impossible in most treatment planning systems. For this reason, in order to investigate the effects of the imperfect linac properties, it became necessary for this research to create a dose calculation engine that could modify or accept modified treatment plan definitions.

In order to assess the magnitude of systematic errors on the linacs at Addenbrooke's so that changes in dose delivery could be studied, extra software was required. High-precision measurements of MLC leaf positions were needed, and there existed no commercial method that provided an accurate process that was compatible with methods employed at Addenbrooke's. Software was therefore created for this research that enabled the measurement of MLC positional accuracy. This software can be implemented at any site, and the measurement of MLC parameters is critical in understanding how IMRT accuracy can be improved.

## Materials and methods

### MLC measurement software

In measuring the location of MLC leaves and collimating jaws, it is not the absolute position of the collimator that is of interest. Instead, it is the position of the radiation field edge that is defined by the collimating device. The field edge is defined as the level of radiation that is 50% of that on the central axis of the linac (for square, symmetric fields). For other field shapes and off-centre beams, the field edge is defined as 50% of either the centre of the beam aperture or the dose that would be on the linac's central axis if the field were symmetric. Radiation scatter, transmission and absorption through and around the leaves contribute to a penumbra whose centre may not align with the projection of the MLC leaves and therefore measuring the position of the MLC is largely uninformative. The in-head light source of the linac can show the approximate locations of beam collimators, as the location of the light and mirror are designed to align the light source with the radiation source. However, the mirror or light source can easily move, and the light field edge shows the edges of any opaque objects rather than the edge of the radiation field. Round-ended MLC leaves will have a light field that is clearly smaller than the 50% radiation definition as the latter is formed by radiation penetrating the opaque rounded ends.

The most reliable way of measuring the 50% field edge is therefore to measure the radiation itself directly. Any method that requires human interaction with equipment is clearly not practical for radiation protection reasons. Remote detection of the radiation is therefore required. Ionisation chambers could be used to detect field edges, but the equipment required to move the chamber automatically to find the field edge is not always available and positional precision and accuracy of the equipment needs to be externally checked. In addition, ionisation chambers are relatively large compared to the MLC leaf accuracy that is demanded by IMRT and measure radiation

over their whole volume. Determining a radiation field edge is therefore difficult to do with accuracy.

Imaging of the radiation field provides 2-dimensional information that can allow determination of all MLC leaf positions simultaneously. There are two common forms of radiation field imaging for megavoltage X-ray beams: radiographic film and electronic portal imaging devices (EPIDs). EPIDs are useful because they require no developing, can provide an instant digital image and use no consumables. However, the contrast of portal images (PIs) can be low, and the pixel size (at isocentre) is between 0.25 and 2.3 mm (Langmack 2001, Boyer *et al* 1992, Munro 1995). Although sub-pixel interpolation can be used, the general accuracy of any measurement taken from a PI can not reasonably be expected to be better than the pixel size. In addition, EPIDs cannot easily be marked to show the isocentre of the machine and beam axis, which makes measurements of absolute leaf position difficult. Although positioning of EPIDs is improving, the calibration can drift over time and thus measurements of the MLC positions would drift with it. Radio-opaque markers can be placed in the radiation beam, but in order to be easily visible on the image, may need to be several pixels in size, which may make finding their exact location difficult. In addition, EPIDs are relatively fragile pieces of equipment with a limited range of possible positions.

Radiographic film, on the other hand, has excellent properties for measuring radiation beams. Sheets of film can be placed anywhere in the beam, at any angle, and are easy to handle, many coming in ready-pack envelopes. The possible resolution of film is also excellent, with films such as X-Omat V[1] having a grain size of approximately 1.5 μm (Haus *et al* 1997, Cheng and Das 1996), while EDR2[1] film has grains around 1/10 this size (Olch 2002). Contrast is generally good with radiographic film, and can be enhanced for particular applications with the addition of metal plates or build-up material (Plane 1985). Any practical way of introducing radio-opaque markers

[1] Kodak Eastman, Rochester, NY

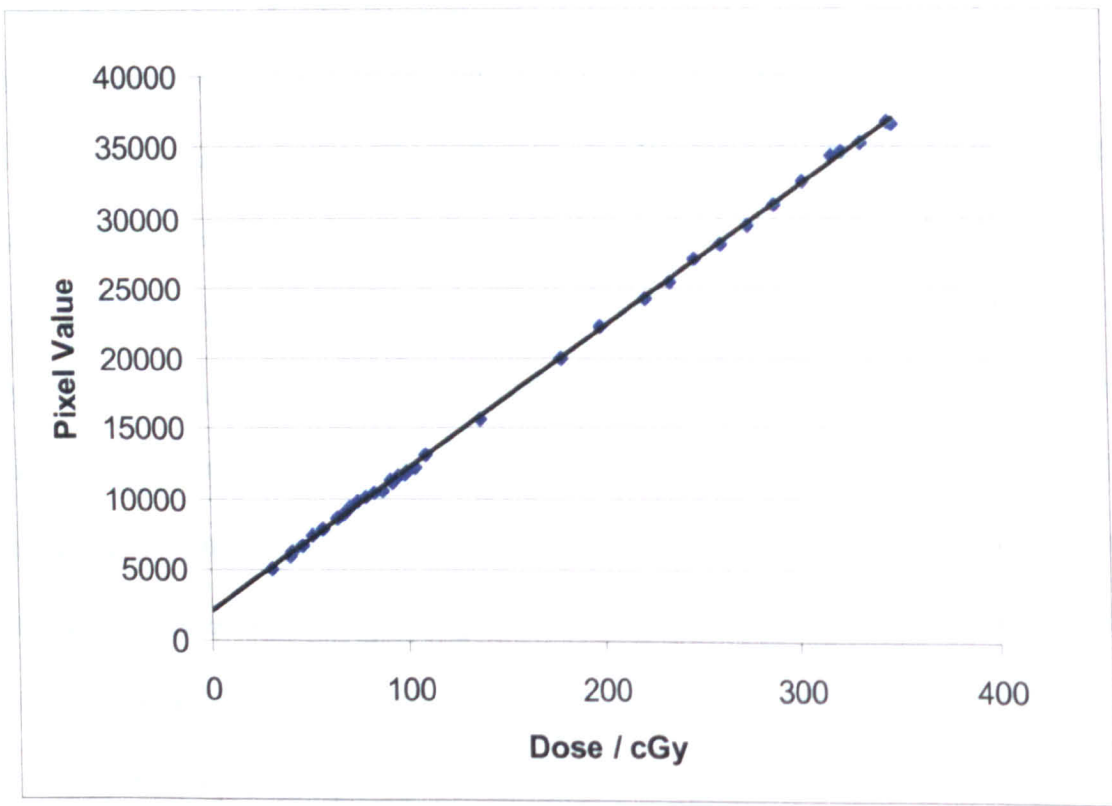



**Figure 4.3:** Graph showing the optical density response of EDR2 film with absorbed dose. The vertical scale shows pixel value as obtained from a VIDAR film digitizer. The black line shows an excellent linear fit to the data points.

onto PIs can be equally applied to film while film can also be physically marked to give location information relative to the linac. The easiest method of marking a film in such a way is by using pinprick holes in the film and envelope. This causes a large (approximately 1 cm diameter) circle of darkened film where light enters the light-tight envelope, but also makes a tiny hole in the film itself. The large darkened circle provides an easy way of quickly locating the pin prick, and if the pin pricks are placed away from areas of interest, they need not cause a problem with measurement of specific features. Features measured on a film also directly correlate with their real size, whereas EPIDs require pixel-size calibration that may not be accurate and will likely require checking on a regular basis.

Accurate, precise and objective measurement of features on radiographic film is most conveniently achieved using digital methods and software. Film therefore needs to be digitized, which can be accomplished using readily available scanners. Using such a device, a film can be digitized to as high a resolution as 300 pixels per inch (dpi). A standard PC desktop scanner could be used, achieving resolutions as high as 4800 dpi, although the spatial and optical accuracy of such scanners may not be as reliable as dedicated film digitizers. Modern verification films, such as EDR2 are designed to have a linear optical density response to absorbed dose through most of their useful exposure range, which is achieved with a fine, uniform grain size. Film digitizers can convert transmitted light to optical density and thus pixel values in digital images can correspond linearly to dose, enabling easy location of the 50% field edge (see figure 4.3). If non-linear film is used, or a film digitizer does not provide linear pixel values with optical density, then a calibration graph can be created to equate dose to pixel value.

Once an exposed film has been digitized, features in the image can be automatically found with high precision. As with portal images, interpolation can be

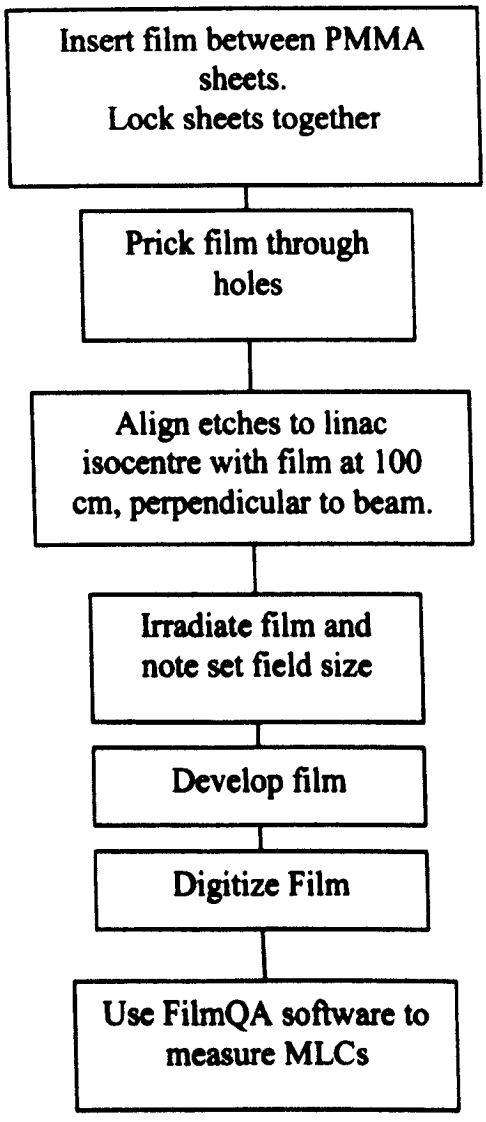



**Figure 4.4:** Flowchart showing the process of measuring MLC positions using film and the FilmQA software

used to acquire sub-pixel measurements, although the absolute accuracy can only be taken to be the pixel size itself, as low as 0.085 mm for 300 dpi resolution. Software was written by the author for this thesis which can automatically find field edges and therefore MLC leaf positions in digitized images of radiographic film. The software, called FilmQA, reads an image file from the film digitizer and calibrates pixel sizes to physical distance using the scan resolution information in the file header. If objects in the image are a known distance apart, for example pin prick holes, then these can be identified with mouse clicks to check the accuracy of the spatial calibration. This can eliminate inaccuracies in the quoted scan resolution of the digitizer and also the effects of film stretching or shrinking during processing, and the calibration can be easily carried out in several directions on the image. The FilmQA software interface provides zoom and contrast tools which enable the user to manually identify small features in the image. If pin pricks are made in the film along the cardinal axes of the linac coordinate system, these can be located and the image can then be automatically rotated to square alignment and the isocentre identified. The software then measures the optical density outside of the irradiated area and at the centre, and identifies the field edges and MLC leaf positions based on the 50% irradiation level. The user can input a different edge detection level if the film response is not linear with dose or for finding 20-80% penumbral widths. The positions of the identified features are saved to a spreadsheet file for subsequent analysis.

To assist with irradiation of films, a PMMA film jig was made. The method of operation of the MLC measurement process is shown in the flow diagram figure 4.4. The jig consists of two sheets of PMMA, each 10 mm thick which lock together snugly using a series of holes and pegs such that the two sheets cannot move laterally once placed together. The sheet of film is placed between the sheets, ensuring that it is kept flat during the irradiation. Along the top surface of the jig are etched two perpendicular

lines, which cross at the centre of the sheet. These lines allow fast and easy setup to the linac room lasers or light-field reticule markings. Along these lines, small (<1 mm) holes are drilled through the top sheet to allow a pin to mark the film once in place. The locations and separations of these holes are known precisely as the jig was manufactured using computerised equipment in our workshop. The accuracy of the hole location was checked with several rulers, and was found the be less than measurable (±0.5 mm) over a distance of 400 mm, or <0.1% variation. Several sets of holes are provided to enable films of different sizes to be pricked near their edges. The whole assembly can have extra build-up material placed in front or behind it if the radiation dose to the film is insufficient. Since 10 mm is less than the depth of maximum dose for most megavoltage X-rays, it was been found that adding extra build-up produced a more repeatable optical density on the film.

The setting up of the film irradiation jig to the linac is the most crucial part of the process. Although field widths between pairs of MLC leaves can be measured from any film, individual MLC leaf positions can only be measured from a suitable reference point, which must be determined. Pin pricks on a film, if located along the central axes of the radiation beam (as they are for the PMMA jig above), allow a beam-axis reference point to be defined. By locating two pin pricks along the same line, the FilmQA software can also correct for rotation of the film during setup on the linac and subsequent digitizing. The PMMA jig must be set up with the etched lines corresponding to the cardinal axes of the linac. If patients are usually set up using the in-room laser lines, then it may be appropriate to align the etch lines to the lasers, else the light-field reticule markings can be used.

Once the PMMA jig is aligned to the linac and is positioned so that the film is at isocentre (usually 100 cm from the radiation source) and perpendicular to the beam, the film is irradiated. The size of the field set on the machine should be noted to be

compared with the sizes measured. If required, several fields can be imaged on one piece of film to calibrate leaves at several locations across their range of travel. Any field size can be used, although the most complete information will be gathered from a field with the jaws that are perpendicular to the MLC leaves open as far as possible in order to image the maximum number of MLC leaves

**Creation of an independent dose calculation engine**

In order to investigate the effects of systematic changes in linac behaviour on the dosimetry of treatment plans, it became necessary to create an independent dose calculation engine (ICE). ICE needs to be capable of reading the details of a treatment plan created from an external treatment planning system and calculate the dose distribution created by the arrangement of beams defined within it. ICE also needs to be capable of changing parameters read from the file or dealing with files that contain data precision better than that which the external treatment planning system can deal with. For example, most treatment planning systems and linacs can only set MLC leaf positions in steps of 1 mm. In order to simulate the effects of unwanted MLC motion that are smaller than these, ICE is required, as the treatment planning system will likely round or truncate non-integer numbers of millimetres. ICE also needs to allow for variable machine parameters, such as the input of non-linear response of dose to monitor units, or behaviour that is a function of machine state, such as MLC motion under gravity as a function of gantry angle.

The MATLAB programming environment provided the necessary tools for the importing of patient treatment plans as produced by commercially available treatment planning systems. The RTPConnect[1] file format was chosen, as it is supported by many treatment planning systems, and while containing all the information of more complicated files such as DICOM-RT, is a simple plain-text file format that is easy to

[1] Impac Medical Systems, Mountain View, California

read and easy to modify external to ICE. In addition, the DICOM-RT specifications allow several optional entries in files which make importing and parsing these files more difficult for files from differing systems. The RTPConnect format, while also having optional fields, has a fixed format with blank entries for all non-included entries.

A dose calculation engine has therefore been written in the MATLAB language to calculate the dose distribution created within a patient or phantom of arbitrary shape. The engine can calculate the dose distribution based on the parameters specified by the treatment plan file, or based on modified versions of those parameters. The program can work to arbitrary precision since it does not need to export to any external systems.

Commercial vendors of treatment planning systems spend many thousands of pounds and thousands of man-hours perfecting the dose calculation engines within their systems. That level of development was impossible within the scope of this research, so a simplified dose engine was decided upon from an early stage. Accurate modelling of radiation beam characteristics is the most vital step in creating an accurate dose calculation engine. Beam characteristics change with field size and shape, measurement depth, initial beam energy, position across the plane of the beam's profile, distance from the source of radiation and the material through which the beam passes. All of these changes in beam properties need to be modelled accurately to give absolute dose calculations that are correct. However, the dose calculation engine to be created for this work only needs to be accurate enough to show relative changes in dose distributions as a function of changing machine parameters. For small changes of parameters such as dose linearity with MUs and small changes in MLC leaf positions, less accurate beam modelling is required.

The dose calculation engine extends a simple 0-dimensional (point-based) calculation technique into three dimensions. For each beamlet, the engine computes the depth of the calculation point, the distance of the point from the radiation source, and

the distance of the point from the beamlet's edges when the beam aperture shape has been projected from the radiation source to the point. The beam profile is assumed to be flat across the open aperture with an edge penumbra the same as that measured for 10 cm × 10 cm fields. Thus the initial fluence is calculated by finding the relevant point on the penumbra curve for the distance from the beamlet's edge and multiplying by the number of MUs for the beamlet. The fluence is then multiplied by an attenuation factor based on the source distance (to account for the inverse square law) and the depth of patient or phantom the beam has passed through. Scatter of radiation both in the head of the linac and in the patient mean that larger fields give more radiation per MU than do smaller fields. Thus the fluence is then multiplied by an output factor for the field shape and size. An equivalent square size of the field is calculated and the output factor is found from a lookup table.

When all of these factors are multiplied together, the dose at a point from a single beamlet is found. The total dose at the point is found by repeating the calculation for all beamlets in a treatment plan. This dose derivation method is accurate to the order of approximately 3% for absolute calculation of dose on the central axis of the beam for the point at the isocentre of the linac coordinates, and is often used to check the output of treatment plans from treatment planning systems. The dose engine developed here calculates the dose across specified 2-dimensional planes by simply repeating the calculation at every desired point. The extension of the method from a point to a 2-dimensional dose distribution is less accurate, due to the linac beam not being flat across its profile and scatter properties of patient contours that are not perpendicular to the beam direction. However, the absolute dose distribution is not required for this work, only the relative changes when small adjustments are made to the treatment plan.

The dose engine, then, can calculate the dose distribution created with imperfect machine parameters. A lookup table of cGy / MU derived from the measurements carried out below simulates dose errors as a function of the number of MUs used for each beamlet. A simple MLC-shift function also moves the MLC leaf positions in the treatment plan based on the gantry and collimator angles – if the two angles combine to exert gravitational forces on the leaves, the leaves are moved by a variable, specified amount. Finally, random errors may be introduced, such as changes in MLC leaf positions within the manufacturer's quoted tolerances and small (<1%) variations in dose output due to day-to-day variability.

**IMaker for fluence map re-segmentation**

The algorithm used to segment an intensity modulated fluence profile may affect the magnitude and locations of dosimetric errors caused by systematic changes in linac behaviour. In order to be able to segment sample treatment plans in different ways, an external program, ImFAST[1] was used. ImFAST reads specially-formatted files that contain a description of the fluence maps used for a treatment, and exports RTPConnect files with the segmented beam arrangements. For some of the sample plans used, it was necessary to recreate the fluence profiles from plans that had already been segmented such that the fluence maps could be segmented in other ways. Software was therefore written to read and parse the information in RTPConnect files and export a fluence map by summing the areas of open fields for beam segments contained within the files. This then allowed ImFAST to re-segment the plans using a variety of built-in segmentation algorithms. These algorithms are: Slice, Standard, Platform, Standard Optimal and Platform Optimal (Siochi 1999, Bortfeld *et al* 1994). The Standard method is what is referred to as 'sliding window' in chapter 1 and the Slice method is the 'close down' method described. Platform is a hybrid method which treats a large field to the level of

[1] Siemens Medical Solutions, Concord, California

the minimum value in the fluence map and then delivers the remaining modulated fluence using the Standard method. Platform is useful for fluence maps where the overall level of fluence is high with little modulation. The Standard Optimal method tries several configurations of segments to find the one with the minimum treatment time. The Platform Optimal method uses the same method as Platform but uses Standard Optimal for the non-base part of the fluence map.

Both the sliding window and close-down method produce fluence maps with equal steps between the different levels of fluence in the map, due to the way that the fluence is divided (see figure 1.10). Therefore all of the numbers of monitor units for each segment for one particular beam should be the same. Division of a fluence map into individual segments allows recreation of the fluence through simple addition. However, the radiation output from a linac is less than 1 cGy/MU for small fields and larger than 1 for larger fields. This is due to the radiation source not being a point, but a diffuse area, spread out by the flattening filter used. Small fields 'see' less of this source than do large fields, and thus the radiation output is lower than 1. It is common to define the output as unity for fields of $10\times10$ cm$^2$. Treatment planning systems such as Xio[1] therefore changes the numbers of monitor units obtained from the simple segmentation to correct for the different output factors and raise the level of radiation coming from small fields. This means that segments have slightly different numbers of monitor units and thus a recreation of the fluence map will be incorrect. The software created for this project 'un-corrects' the numbers of monitor units by using the assumption that the steps between fluence map pixels should be the same. The software automatically determines the number of fluence levels by a histogram binning method. The approximate field size is then calculated for each segment and if it is less than 10x10 cm$^2$, the number of MUs is rounded down to the nearest level. The software then exports the fluence map in a format that ImFAST can read.

[1] CMS, St Louis, Missouri

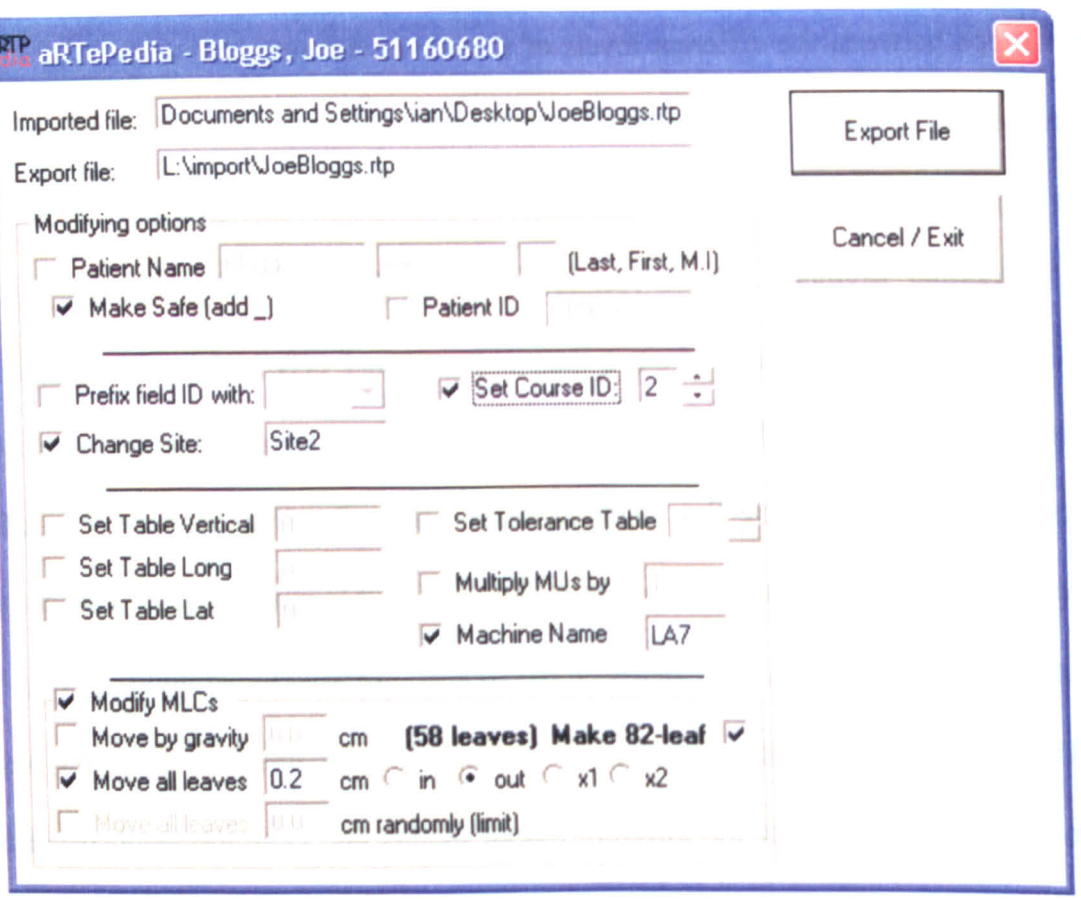



**Figure 4.5:** Interface of RTPedia software. An RTPConnect file is imported and the data fields are populated from it. These data can then be modified and exported as a new file.

The software is designed to work with RTPConnect files generated by the Xio or Corvus treatment planning systems. Both systems can be set to create fluence maps with differing resolutions in the MLC leaf travel direction. Addenbrooke's uses settings of no smaller than 1 cm to prevent the segmentation creating very small fields, and thus IMaker creates fluence maps with $1 \times 1$ cm$^2$ resolution, equal to the resolution of the original fluence maps.

**RTPedia**

Verify and Record (V&R) systems exist as an interface between treatment planning and the linac delivery of radiation, and are essential for delivering complex treatment such as IMRT. V&R systems store complex field and MLC configurations and automatically download them to the linac systems for easy and reliable treatment. However, because V&R systems are designed as a clinical system, they implement various features to prevent accidental (or deliberate) irradiation of patients with the wrong treatments. One such restriction is that all treatment fields imported into the V&R system need to be properly divided into a hierarchy of treatment site, treatment course and unique field identification numbers. As such, it is not possible to import the same treatment plan, or even the same treatment plan that has been segmented differently, as the newly imported fields will overwrite the old ones. Editing RTPConnect files is possible, but for files that contain many beams and IMRT segments, this is laborious; RTPConnect files also contain cyclic redundancy check (CRC) numbers to check that no data has been modified. Thus an RTPConnect file cannot be intentionally modified without recalculating the CRC number.

A software program, called RTPedia, was therefore created for this project, which is capable of modifying RTPConnect files in a number of ways. Figure 4.5 shows the interface of RTPedia and the range of entries in the file that can be modified.

There are several simple, useful changes, for example prefixing a field ID with a letter to allow the same treatment plan to be imported into the V&R system twice, or changing the treatment course number or site text to the same end.

At Addenbrooke's there are 5 IMRT-capable Siemens linacs (as of 2005), with slightly different configurations. Three machines are Primus-class and two are Oncor-class. The Oncor machines have opposing banks of 41 1-cm-wide leaves, whereas the Primus machines have banks of 27 1-cm leaves with 6.5 cm leaves at the ends to make the maximum field size up to 40 cm. The central 27 cm of the Primus and Oncor machines are identical in behaviour and thus treatment plans that use only this area are transferable between machines. However, the V&R system in use (Lantis, from Impac/Siemens) does not allow plans to be imported onto the 'wrong' machine. Thus a useful feature of RTPedia is to be able to change the machine name defined for the treatments, or to change the number of MLC leaves in the plan from 58 to 82 or vice versa. Converting from 58 leaves to 82 is trivial, as the outer sets of extra leaves are set to the same as the 6.5 cm leaves from the 58-leaf configuration. Converting from 82 to 58 leaves is more complex. If the extra leaves of the 82-leaf configuration are used, then they cannot be adequately simulated with one 6.5 cm leaf. In this case, RTPedia warns the user that the new treatment field shape will be different, and sets the 6.5 cm leaf position to that of the innermost leaf that is equivalent (leaf 7 or 35). Lantis does provide the option for changing the treatment machine once a treatment plan has been imported, and converts the leaves in the same way as RTPedia. However, Lantis does not give a specific warning that the field shape will change.

When the user has selected the changes to be made to the RTPConnect file, RTPedia exports the file, including recalculated values of the CRC numbers so that the new file is valid. As a precaution against accidentally importing modified files over real

patient entries in the V&R system, an underscore character can be added to the patient name such that no match is found in the clinical database.

In order to validate the errors that ICE would be calculating, treatment plans were required that had MLC leaves moved in the way that ICE was simulating. Even on an imperfect linac, if a treatment plan is delivered for which all of the MLC leaves have been deliberately moved out by 1 mm, the relative dose difference should be approximately equal to that calculated by ICE for the difference between a perfect machine and a treatment plan with all MLC leaves moved. Thus additional features were added to RTPedia to enable treatment plan files to be modified in such a way. RTPedia is capable of moving all MLC leaves into the field or out of the field in integer millimetre steps. Although steps smaller than these are required by ICE for MLC modelling, it is able to move the MLCs within itself. RTPedia is only needed for modifying treatment plans that will eventually be treated on a linac, and the control software of the linacs only supports millimetre steps. RTPedia runs simple checks on the MLC leaf positions and cleans up any irregularities. For example, a pair of closed leaves cannot be moved into the field any further, and so RTPedia does not move these leaves. Moving leaves out of the field is easier. Although it has been seen that closed MLC leaf pairs do not open by the same amount as open MLC leaves move when the gantry is rotated, RTPedia still opens these leaves slightly. The segmentation algorithms used for IMRT at Addenbrooke's use a 'single aperture correction' which ensures that no multiple openings are created in any IMRT segments and thus no MLC leaf pairs are closed in the centre of beam areas. Thus all closed MLC pairs are outside of the treatment area and are therefore blocked by the collimating jaws perpendicular to the MLC leaves, which automatically move into the edges of the open MLC leaves.

**Validation of software**

Before use, the various pieces of software written for this project have to be tested and validated both as standalone applications and interfacing with third-party products. The software that needed testing is: FilmQA, ICE, RTPedia and IMaker.

RTPedia was validated through careful inspection of files after export from the software. Treatment plans were imported and exported with no change and the files were identical. Changes to the files were largely validated by importing into the V&R system Lantis, which has stringent checks on incoming files. Specific changes to the files were also checked by inspecting the raw data in the text files. Under all circumstances, RTPedia behaved as designed.

IMaker was checked by importing known RTPConnect files to ensure that the system could deal with simple files. More complicated files were imported from real treatment plans developed in Xio and Corvus. The resulting fluence maps were compared with fluence maps recreated from the files by hand. Rounding of the numbers of monitor units was found to be correct in all cases, aided by the integer-only export of both Corvus and Xio. Fluence maps were then passed to ImFAST for segmentation. The segmented RTP files were then put back into IMaker to check that the same fluence maps emerged. The segmentation algorithms within ImFAST modified numbers of MUs in the files to account for field size outputs in the same way as Xio and Corvus did originally. Multiple cycles of fluence map creation, segmentation and map recreation were performed, with no changes to the fluence maps observed.

FilmQA was checked for physical accuracy by opening generated images containing objects of known dimensions and with known rotations. FilmQA was able

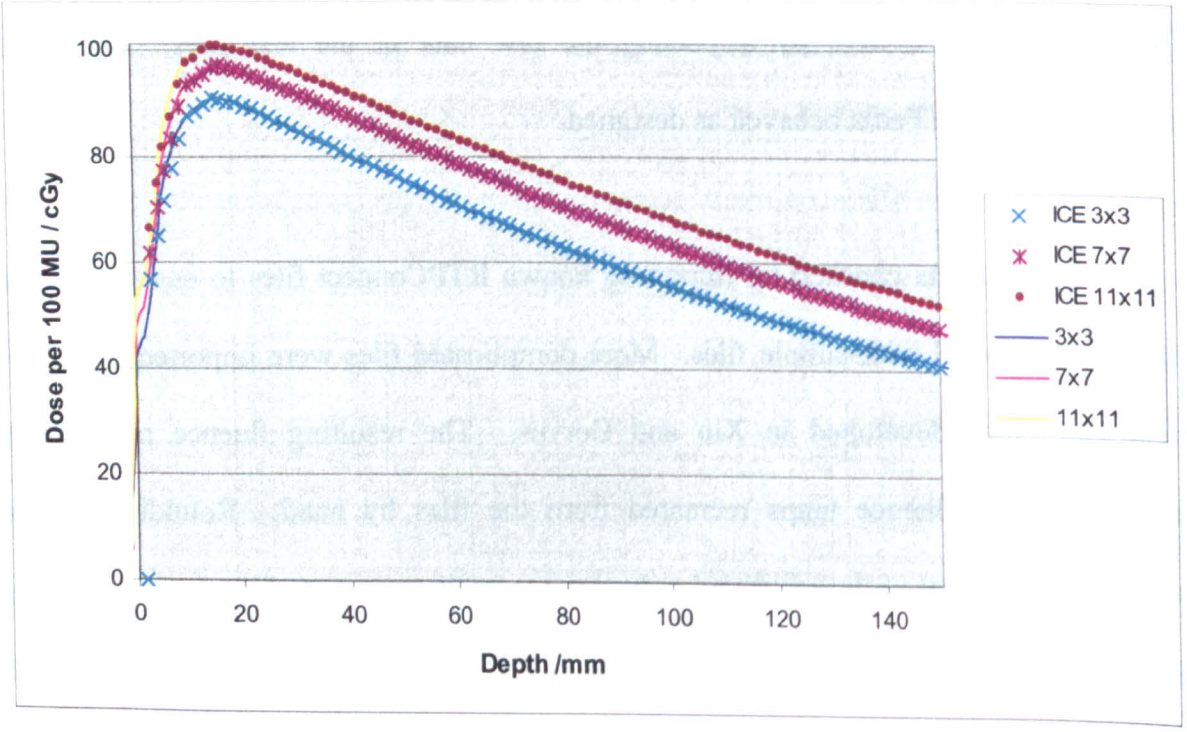



**Figure 4.6:** Depth-dose profiles for 3×3, 7×7 and 11×11 $cm^2$ field sizes at 100 cm SSD. The solid lines show the values measured with an ionisation chamber and used as gold-standard data at Addenbrooke's. The data points show values calculated by ICE.

to automatically rotate images back to their original orientations and measured the distances between pixels to arbitrary precision. Reading of pixel values from images was confirmed by importing images created from within the MATLAB environment as well as from Paint[1]. Detection of specified pixel values along profiles was tested by creating images with known objects and checking that all objects were successfully detected. Saved data files were opened in Microsoft Excel to validate the data writing process, and all data values were compared against the original images and raw values from within the MATLAB environment.

ICE underwent significant levels of testing during its creation as any large piece of software does. Most functions within the program were tested as standalone components, and because they are only called from within ICE, the inputs and outputs can be closely controlled. The only external interfaces of ICE are reading data from RTPConnect files and CT images. Both are well-structured file formats, and simple checks in the ICE program warn the user if the incoming data is detected as invalid. Dose distributions were calculated and were verified as qualitatively similar to dose distributions calculated for the same treatments in treatment planning systems such as Xio and Corvus. Point doses at the isocentre for two clinical IMRT plans were calculated and agreed with the treatment planning system used to develop the plans to within 5%. Simple tests such as the generation of depth-dose profiles and cross-beam profiles also showed the accuracy of the system to be good for simple beam arrangements. Figure 4.6 shows depth-dose profiles calculated for squares fields of size 3, 7 and 11 cm. The agreement between ICE and measured data is excellent, with differences of less than 0.5% for all depths greater than 2 cm, and a DTA of less than 1 mm for depths less than 2 cm. Calculation of dose from multiple beams is simple addition of that from individual beams, which was checked with a set of identical beams

[1] Microsoft Corporation, Redmond, WA

from different gantry angles. Modification of machine parameters was checked with debug output of internal variables before and after modification during the execution of the program. Integer-millimetre changes were made to treatment files and recalculated in Xio and the dose differences found in ICE for sub-millimetre changes were between the integer millimetre changes calculated by Xio.

The dosimetric accuracy of the calculation of smaller fields, typical of IMRT treatments, was not as good in ICE as for the larger fields. However dose differences between unmodified and modified machine parameters were in good agreement with those from other treatment planning systems in which individual small fields were separated from treatment plans.

The calculation of dose delivered from modified beams was compared to modified beams delivered on linacs. RTPedia was used to create treatment files with integer millimetre movements on all leaves, and these treatments were delivered to a phantom block with an ionisation chamber at the centre, as well as an unmodified treatment file. The dose at the ionisation chamber was calculated using ICE for MLC movements of -2 to +2 mm and the change in dose was found to be linear with MLC movement. This meant that even if the linac was not set up perfectly, the difference between a perfect treatment file and a modified treatment file should be the same as when those two files were used for calculation in ICE. ICE calculated a change in dose of 3.4% per millimetre for a chosen treatment plan, which agreed well with the 3.1% change in dose as measured by the ionisation chamber.

**Treatment plans and treatment planning systems**

In order to investigate the effects of the behaviour measured in real cases, two IMRT patient plans were chosen from those planned at Addenbrooke's as trials during the implementation of an IMRT program. The first, Patient H, is a prostate cancer

patient, whose treatment was planned using the Corvus planning system. The treatment uses 5 beam angles with 5 intensity levels in the fluence modulation. The plan was initially segmented using the ImFast sliding window algorithm, and consists of 50 segments. The beams angles were 0, 60, 120, 216 and 270 degrees, chosen by the treatment planner to give an anterior and a lateral beam for portal imaging purposes and to give treatment beams from all around the body while avoiding direct irradiation of sensitive structures such as the rectum.

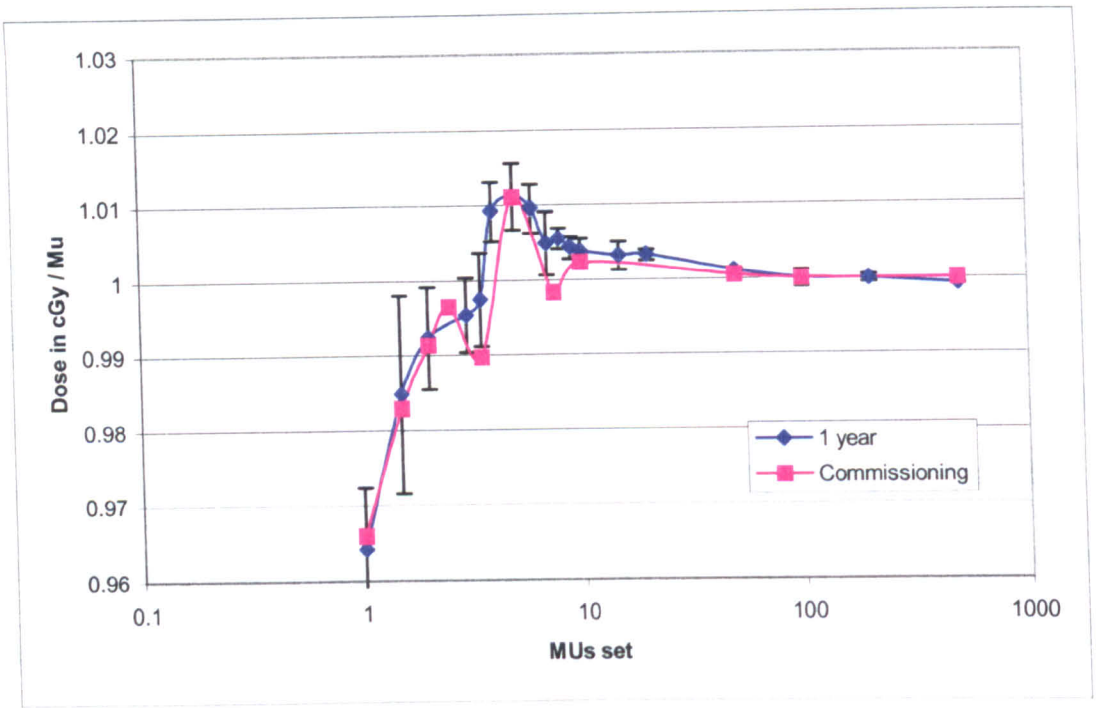
The second patient plan, for Patient S, was a head-and-neck case receiving treatment to the tonsils with attention paid to sparing the parotid glands as part of the PARSPORT trial[1]. Patient S was planned using the Xio planning system, and uses 5 beam angles. These beams are equally spaced around the patient at 0, 72, 144, 216 and 288 degrees. Additional anterior and lateral beams were added for portal imaging, although these were not included in the calculations carried out here as the fields were low in dose and large, open apertures. The treatment plan used 8 intensity levels in the fluence modulation and resulted in 89 segments using the ImFast sliding window algorithm built in to the system.

The treatment planning systems used were Corvus version 4 and Xio version 4.2.0. Both use inverse planning methods in which the treatment planner specifies desired doses to the target volume and sensitive organs, and the planning system optimizes fluence maps to best match those doses. Both systems were set to use 1×1 $cm^2$ pixels in the fluence maps. The segmentation algorithms were set to give apertures no smaller than 2 $cm^2$.

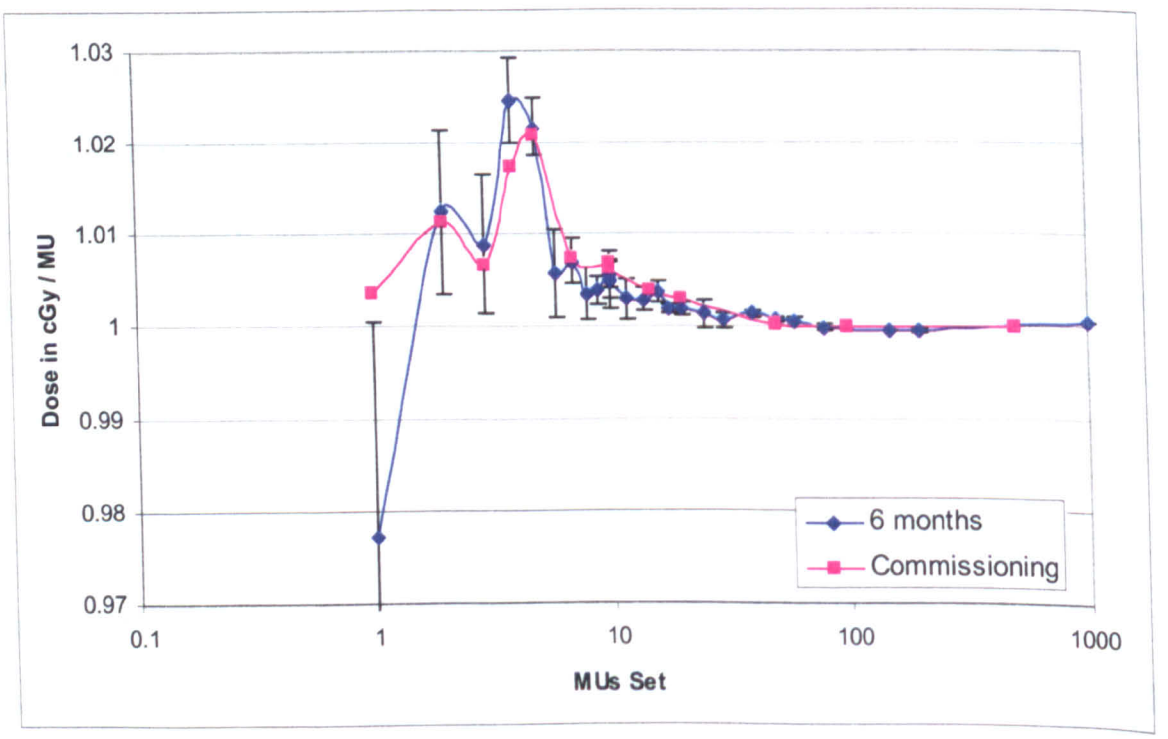
**Experiments**

In order to investigate the effects of clinically relevant systematic errors in linac behaviour, the magnitudes of those errors needs to be established. There are two

[1] Trial contact: Dr Chris Nutting, Royal Marsden Hospital, London. ISRCTN: 48283537




**Figures 4.7 a and 4.7b:** Dose delivered by a linac (LA7 above, LA8 below) at Addenbrooke's as a function of Monitor Units set. The error bars represent the standard deviation of repeated measurements combined with the measurement errors. The pink line shows the data collected at the time of commissioning of the machine, the blue line shows data collected at a later date (as shown in the legend). Errors were not available for the commissioning data. All machines at Addenbrooke's are calibrated to give 1 cGy/MU at 100 MU and the graphs are normalized to that level.



aspects of linac behaviour that were measured, namely MLC motion and positioning, and dose linearity as a function of monitor units set.

Monitor unit linearity was measured in the same setup as linac outputs are measured for ongoing quality assurance checks. A type 30012 Farmer-type[1] ionisation chamber was set up in a block of water equivalent material[2] with the surface 100 cm from the radiation source. The chamber was positioned in a specially-drilled cavity in the block at 5 cm depth. The beam's axis was perpendicular to the surface of the block, which was 20×20×20 $cm^3$ in size. Since IMRT at Addenbrooke's is performed using only 6 MV X-rays, only 6 MV beams were measured, using a field size of 10×10 $cm^2$. A Unidos[1] electrometer measured the charge collected by the chamber. The electrometer was set to measure and beams were delivered 10 times without resetting the counting to improve accuracy of the recorded value for small fields. Outputs were calculated in cGy/MU for the numbers of MU shown in figures 4.7a and 4.7b, corresponding to linacs LA7 and LA8 respectively.

The graph shows wide variability in dose linearity for small numbers of monitor units, which settles to less than 1% for beams above 6 MU. The graph also shows that the machine's behaviour is stable, with the pink line showing data collected at the time of installation and the blue line showing data collected six months or 1 year later. The error bars are calculated based on the measurement precision of the equipment used. Information on how the commissioning data was collected was not available and thus error bars can not be calculated. The largest errors are associated with the smaller fields as the measurement precision of the electrometer used was approximately 5% of the radiation output at 1 MU.

For the purposes of inclusion into ICE, the 6-month/1-year values were used. No IMRT fields at Addenbrooke's use fewer than about 4 MU and the 6-month/1-year

[1] PTW-Freiburg, Freiburg, Germany
[2] WT1 material, St Bartholomew's Hospital, London, UK

data has more points in the 5-10 MU range than the commissioning data, which is where many IMRT segments operate.

MLC behaviour was measured using the FilmQA program and the PMMA jig created to accompany it. Measurements of MLC behaviour were carried out on one Primus and one Oncor machine. The Oncor machines are newer than the Primus machines and have better specifications than the Primus machines for many parameters, such as MLC positioning accuracy of ±1 mm rather than ±2 mm. After the machines]' installation, IMRT was moved from the Primus machines to the Oncor machines, with one Oncor machine being the main IMRT location. Measurements were carried out for this linac and one of the Primus machines for comparison. To compare how the double-focused leaves of Siemens machines fare against the single-focused designs of Elekta and Varian, measurements were performed on Elekta and Varian machines at another hospital[1]. Elekta manufactures only one type of MLC, with 40 pairs of leaves of 1 cm width at the isocentre. Varian manufactures several different MLC configurations, with 52, 80 or 120 leaves. The Varian system tested consisted of 120 leaves – each bank has 40 central pairs of 0.5 cm isocentric width with 10 pairs of 1 cm width on each side to give a total field size of $40\times0.5 + 2\times10\times1 = 40$ cm. The Siemens Oncor linac uses 41 leaf pairs, each 1 cm wide at the isocentre, whereas the Primus machine uses 27 central leaf pairs, each 1 cm in width, flanked by 6.5 cm leaves, to give a total MLC width of 40 cm. Although the Oncor machine has an MLC width of 41 cm, the perpendicular jaws only open to 40 cm. Because Varian and Elekta use an even number of MLC leaf pairs, the central axis is on a join between adjacent pairs. Siemens use an odd number of leaf pairs such that the central axis is in the centre of a leaf.

The tests carried out on the Elekta and Varian machines consisted of film measurements of fields with the jaws perpendicular to the MLC leaves open as far as

[1] Elekta at Royal Marsden Hospital, Sutton. Varian at Royal Marsden Hospital, Chelsea

| Accelerator | MLC Test | X1 leaf variation (mm) | X2 leaf variation (mm) | Field width (mm) | Field vs 'Out at 0' (mm) |
|---|---|---|---|---|---|
| Varian | Moved in @ 0 | ±0.07 | ±0.12 | 103.40 ± 0.17 | *+0.17* |
| | Moved out @ 0 | ±0.07 | ±0.13 | 103.23 ± 0.18 | - |
| | Moved in @ 90 | ±0.11 | ±0.12 | 103.86 ± 0.16 | *+0.53* |
| | Moved out @ 90 | ±0.08 | ±0.06 | 102.74 ± 0.12 | *-0.49* |
| Elekta | Moved in @ 0 | ±0.20 | ±0.17 | 100.19 ± 0.29 | *+0.04* |
| | Moved out @ 0 | ±0.28 | ±0.38 | 100.15 ± 0.24 | - |
| | Moved in @ 90 | ±0.37 | ±0.46 | 100.31 ± 0.25 | *+0.12* |
| | Moved out @ 90 | ±0.52 | ±0.46 | 100.48 ± 0.26 | *+0.29* |
| Siemens Primus | Moved in @ 0 | ±0.61 | ±0.57 | 101.86 ± 0.40 | *+0.31* |
| | Moved out @ 0 | ±0.53 | ±0.51 | 101.55 ± 0.52 | - |
| | Moved in @ 180 | ±0.63 | ±0.66 | 102.90 ± 0.56 | *+1.35* |
| | Moved out @ 180 | ±0.64 | ±0.67 | 103.53 ± 0.67 | *+1.98* |
| | Moved out @ 90 | ±0.51 | ±0.67 | 102.59 ± 0.59 | *+1.04* |
| Siemens Oncor | Moved in @ 0 | ±0.28 | ±0.44 | 100.37 ± 0.48 | *0* |
| | Moved out @ 0 | ±0.28 | ±0.41 | 100.37 ± 0.51 | - |
| | Moved in @ 180 | ±0.40 | ±0.25 | 100.85 ± 0.43 | *+0.48* |
| | Moved out @ 180 | ±0.23 | ±0.44 | 101.05 ± 0.54 | *+0.68* |
| | Moved out @ 90 | ±0.29 | ±0.29 | 100.75 ± 0.44 | *+0.38* |

**Table 4.I:** MLC hysteresis movements as measured on four different linacs. MLC leaves were moved either into or out of the field at the gantry angles shown. The field sizes were set to 100 mm in width and measured using film. All variation numbers are quoted as 1SD.

possible and the MLC leaves set to 10 cm. Films were exposed at gantry angles of 0 and 90 with various MLC histories such as MLC leaves moved in before irradiation, moved out, and then moved in or out prior to gantry rotation to 90 degrees. Films were also taken of the MLC leaves moving in and out at gantry 90.

Tests carried out on the Siemens machines consisted of the same $40\times10$ $cm^2$ field sizes. Films were exposed at gantry 0 with the MLC leaves having moved in, and then with them moved out. Films were repeated at gantry 180. A further test moved the leaves either in or out at gantry 0 and then moved the gantry to 90 degrees.

Table 4.I summarises the measurements recorded for the Varian, Elekta and two Siemens machines. The table shows the average field widths as measured between each pair of leaves, with the standard deviation. Also shown are the standard deviations of the individual leaf banks and the difference between each field width and the 'Moved out at gantry 0' width. Moving the leaves out of the field should provide the smallest field as backlash is not taken up, and the increase in field size therefore represents the maximum possible. For these tests, the X2 bank was defined as the one which is on the lower edge of the field when the gantry is at 90 degrees.

The data in the table shows several interesting comparisons between the machines. The Elekta and Varian machines both have excellent leaf positioning, with MLC leaf separation ('field size') variability of around 0.25 mm and 0.15 mm respectively at all gantry angles. The field sizes themselves are also stable with gantry angle, with the Elekta having a maximum deviation of 0.12 mm and the Varian 0.53 mm between gantry 0 and gantry 90. The Siemens machines show less precise MLC positioning capabilities. The newer Oncor machine has a field size that increases by 0.7 mm between gantry 0 and gantry 180, and the variability in MLC leaf separation is around 0.5 mm. The older Primus machine shows worse stability, with a gantry 180 field nearly 2 mm larger than a gantry 0 field and leaf variability of 0.6 mm. The tests

at gantry 90 show that the field sizes are halfway between those at gantry 0 and gantry 180.

The table shows that the Varian fields were several millimetres larger than the 100 mm that was set on the machine. This is because the leaves on this machine are calibrated by the light field, which depends on the positions of the leaf ends which, being rounded, protrude beyond the radiation field edge. Treatment planning systems that plan for machines with round-ended MLC leaves have a leaf-offset variable which sets the linac to have smaller fields than are required.

**Dose distributions**

Dose distributions were calculated for both patient plans with several variables, investigated individually. Firstly, each patient's treatment plans were calculated in ICE with varying levels of MLC mis-calibration. The MLC leaves were set to all be moved into the field by values up to 2 mm and also out of the field by up to 2 mm. Secondly, the treatments were re-calculated assuming that the MLC banks move under gravity as a function of gantry angle. All of the collimator angles used in the treatment planning were 0°, such that the MLC leaf motion was perpendicular to the gantry rotation axis. This provided the maximum amount of MLC leaf motion under gravity. The MLC motion was applied only to the lower bank of leaves, or both banks near gantry 180. In practice, the critical angle for the motion was not needed as the separation of beam angles mean that there were no beams within 20 degrees of gantry 0 or gantry 180 (except for gantry 0 itself). The amount of MLC gravity motion was varied from 0 to 2 mm. Included in this range were the actual values of MLC motion as measured for the Primus and Oncor Siemens machines in use at Addenbrooke's.

Each patient's treatment plans were then re-segmented using IMaker and ImFAST using the 5 different algorithms available within ImFAST. The dose

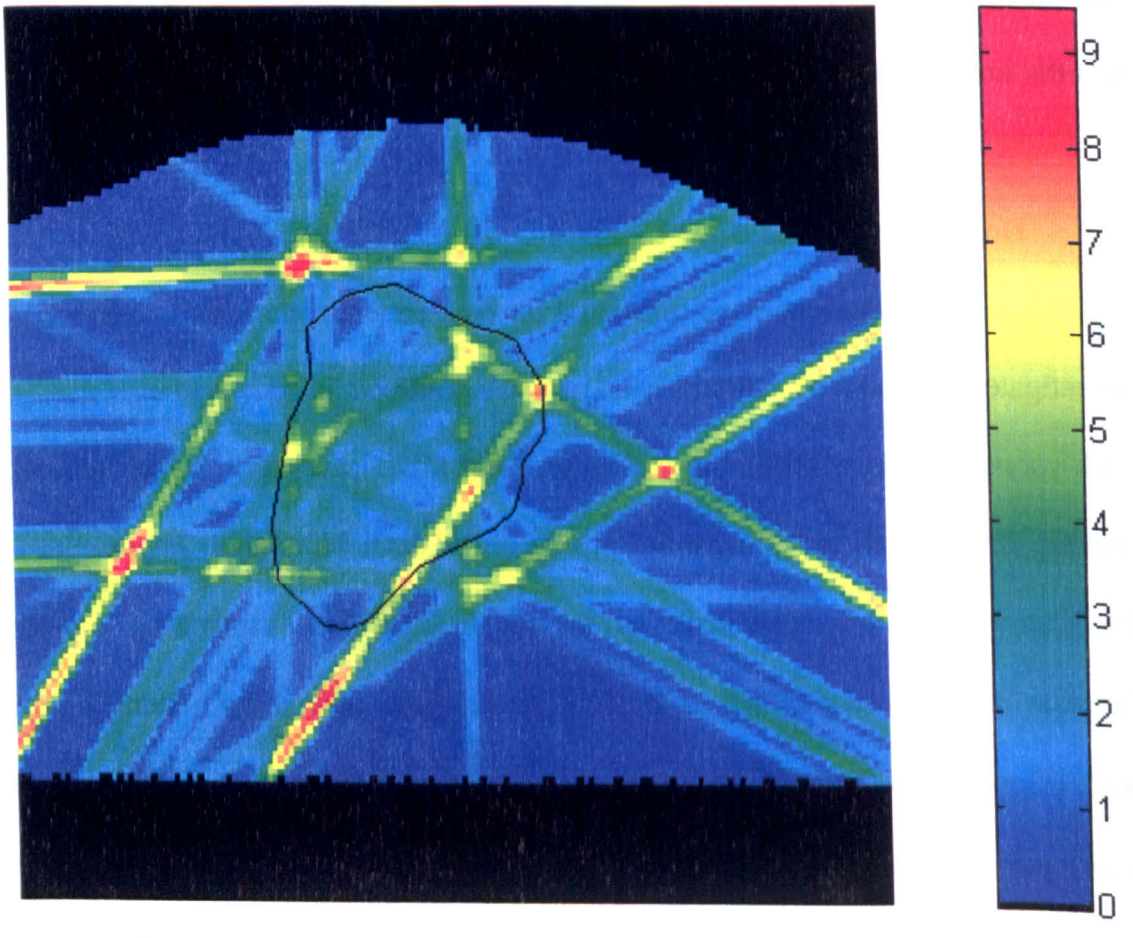



**Figure 4.8:** Dose difference between a treatment with a perfect machine and one with all MLC leaves moved out of the field by 1 mm. The scale shows difference as a percentage of the isocentre dose for Patient H's plan. The dose is calculated for the axial isocentric slice, and the black outline shows the PTV in this slice. Black areas of the plot are outside of the patient.

distributions were calculated with 1 mm gravity motion and then with 1 mm MLC motion out of the fields to investigate if some segmentation algorithms give bigger dosimetric errors due to the effects seen here. For all other tests not focussing on segmentation, the Slice algorithm was used.

Treatments were then calculated with a deliberate MLC offset. An MLC movement into the field equal to half of the gravity motion was introduced. In this way, fields at gantry 0 would be slightly too small, fields at gantry 180 slightly too large. The calculations were carried out to see if a deliberate mis-calibration in MLC position could counter the gravity motion effects.

Finally, dose linearity effects were studied by adding the variable cGy/MU values as measured to plans on perfect machines. The non-linear dose was then added to plans for which the MLCs were allowed to move under gravity to the levels measured for the Siemens machines. This situation should model the dosimetric errors that could be introduced on Siemens machines as used at Addenbrooke's.

## Results

### MLC leaf positioning

Figure 4.8 shows the dose difference, for Patient H's plan, between the plan delivered with perfect machine parameters and with every MLC leaf moved out of the field by 1 mm. Also shown on the plot is the outline of the PTV. The dose distribution was calculated in the axial plane at the isocentre. Clearly visible on the plot are thick lines of overdosing that show the beam directions of 0, 60, 120, 216 and 270 degrees. Hotspots as high as 9% are created where these lines overlap, with the size of 7% hotspots reaching up to 7 mm in diameter.

Figure 4.9 (over page) shows a plot of the Kappa index (see chapter 3) across the plane of dose calculated for figure 4.8. The Kappa tolerances were 3% and 3mm.

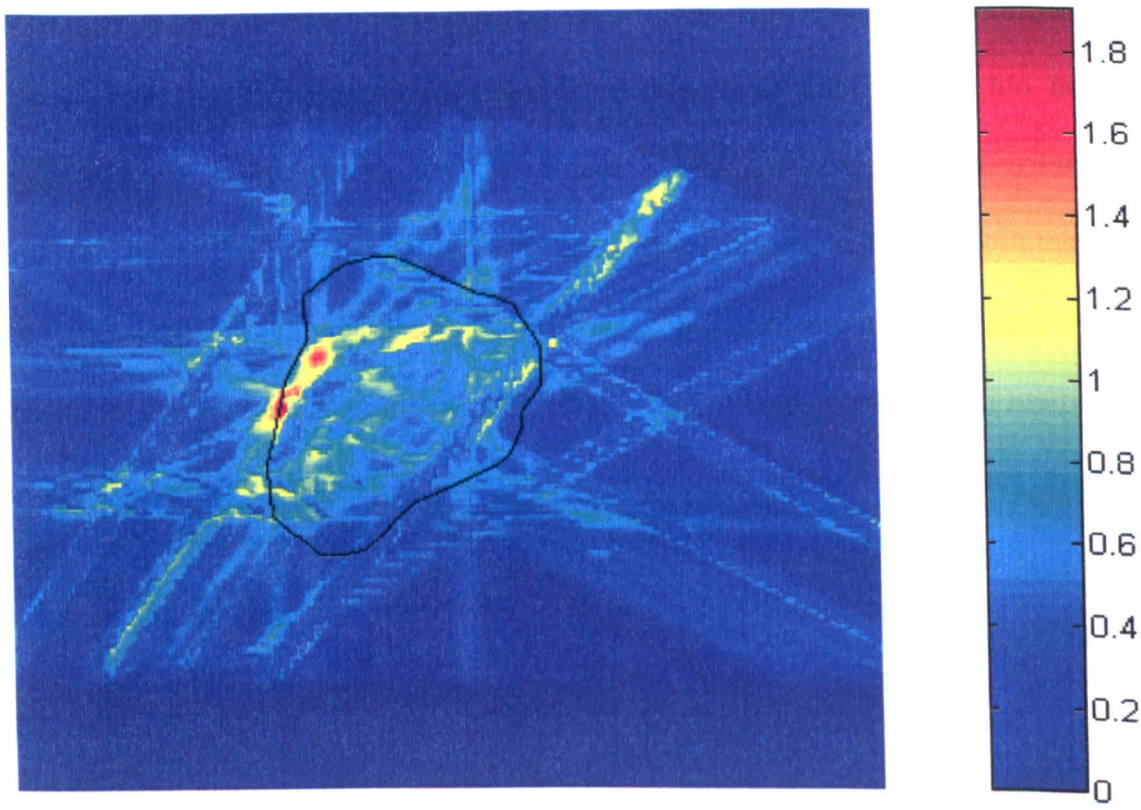



**Figure 4.9:** Plot showing the Kappa index analysis of the dose difference shown in figure 4.7. The Kappa parameters were 3% and 3 mm tolerance. As with figure 4.7, the black outline shows the PTV. A Kappa index of greater than 1 indicates failure of both distance and dose tolerances.

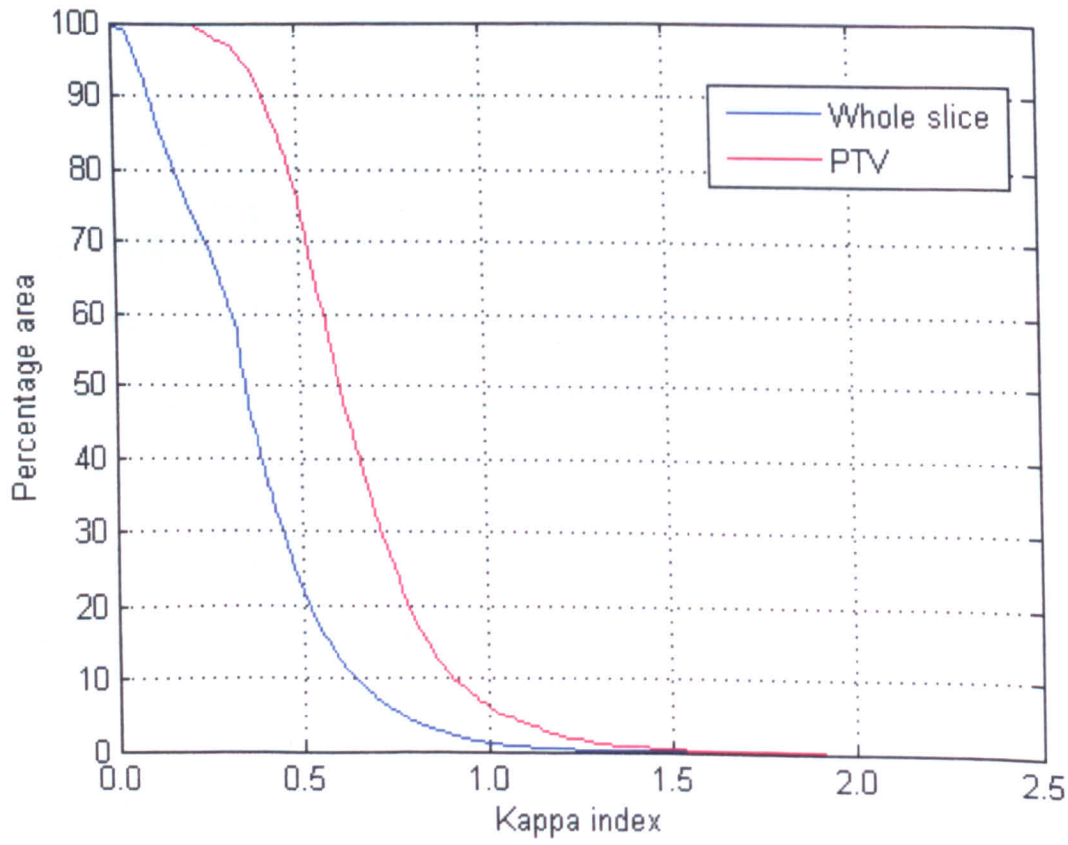



**Figure 4.10:** Cumulative Kappa-area histogram showing the distribution of Kappa index values from figure 4.8. 'Whole slice' refers to the area of calculation shown in 4.8 and 4.9.

Many of the areas that showed large dose differences in figure 4.8 have a low Kappa index because they are near the edges of fields and therefore in high dose gradient areas. However, there are some areas of Kappa greater than 1, most of which are inside the black PTV outline. Indeed, the dose difference plot figure 4.8 shows that many of the largest discrepancies are outside of the PTV, but figure 4.9 shows that the smaller discrepancies in the low-gradient PTV area lead to greater value of the Kappa index than do the larger dose differences outside of the PTV.

Figure 4.10 shows a Kappa-area cumulative histogram for the plot shown in figure 4.9. The histogram shows the percentage area of the whole calculation area (title as 'whole slice') and that of the PTV that is above a certain Kappa index value. Thus 90% of the PTV has a Kappa index value above about 0.4 and neither the PTV nor the body have Kappa index values greater than 1.9. The histogram shows that for this calculated slice, 6% of the PTV structure area exceeds a Kappa index of 1, showing that it is within neither 3% of the correct dose nor 3 mm of the same dose.

The dose differences between Patient H's perfect plan and the case with all MLC leaves moved out of the field by 2mm are approximately double those caused by a 1 mm shift. Figure 4.11 (over page) shows the Kappa-area histogram for this case. The two curves for body and PTV have been moved to the right by a factor of approximately 2. Because of the steepness of the Kappa-area curve this has changed the percentage of the PTV area that has a Kappa index of over 1 from 6% to 74%. The percentage of the body area that has a Kappa index greater than 1 has increased from 1.5 % for a 1 mm MLC leaf movement to 20% for a 2 mm movement. Figure 4.12 (over page) shows the percentage areas that exceed a Kappa index of 1 for MLC movements of -2 to 2 mm, with negative movements indicating a movement into the field.

Figure 4.13 (page 99) shows Kappa-area histograms for Patient S's treatment plan with varying MLC movements, as for Patient H's plan above. The graph

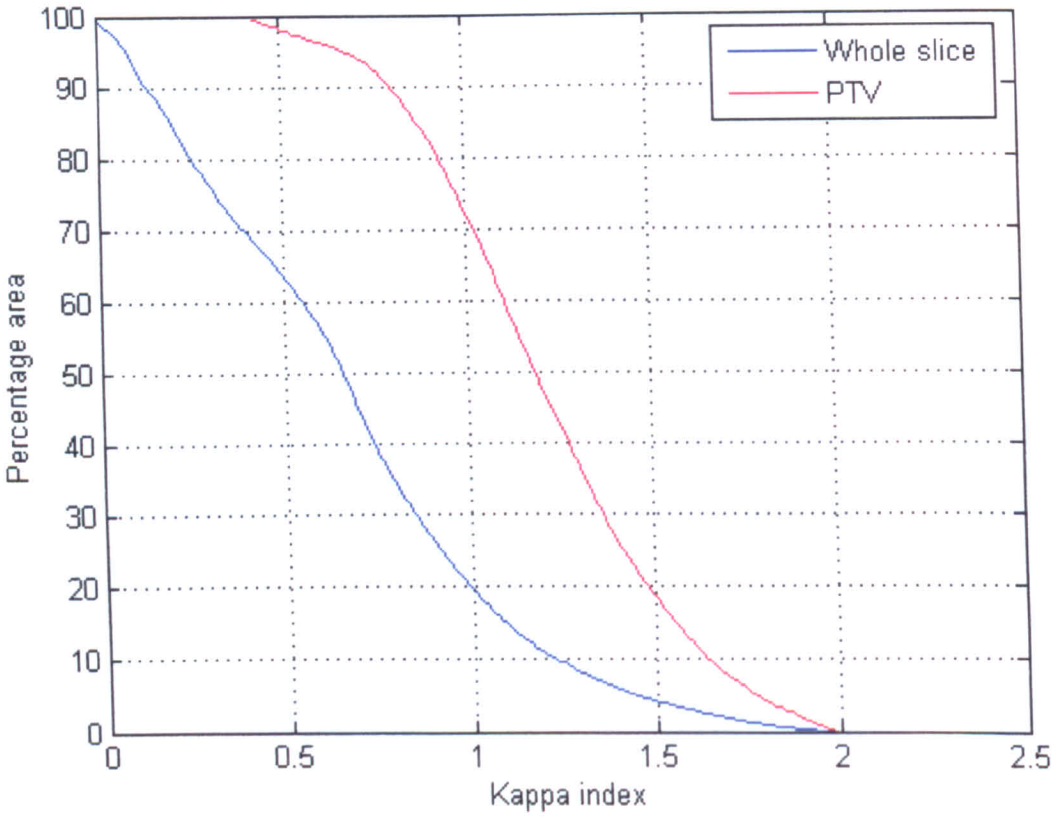



**Figure 4.11:** Kappa area histogram for the comparison between Patient H's perfect treatment and delivery of the treatment with all MLC leaves moved out of the field by 2 mm. The change in parameters is therefore double that for the histogram shown in figure 4.10. Tolerance parameters were 3% and 3mm.

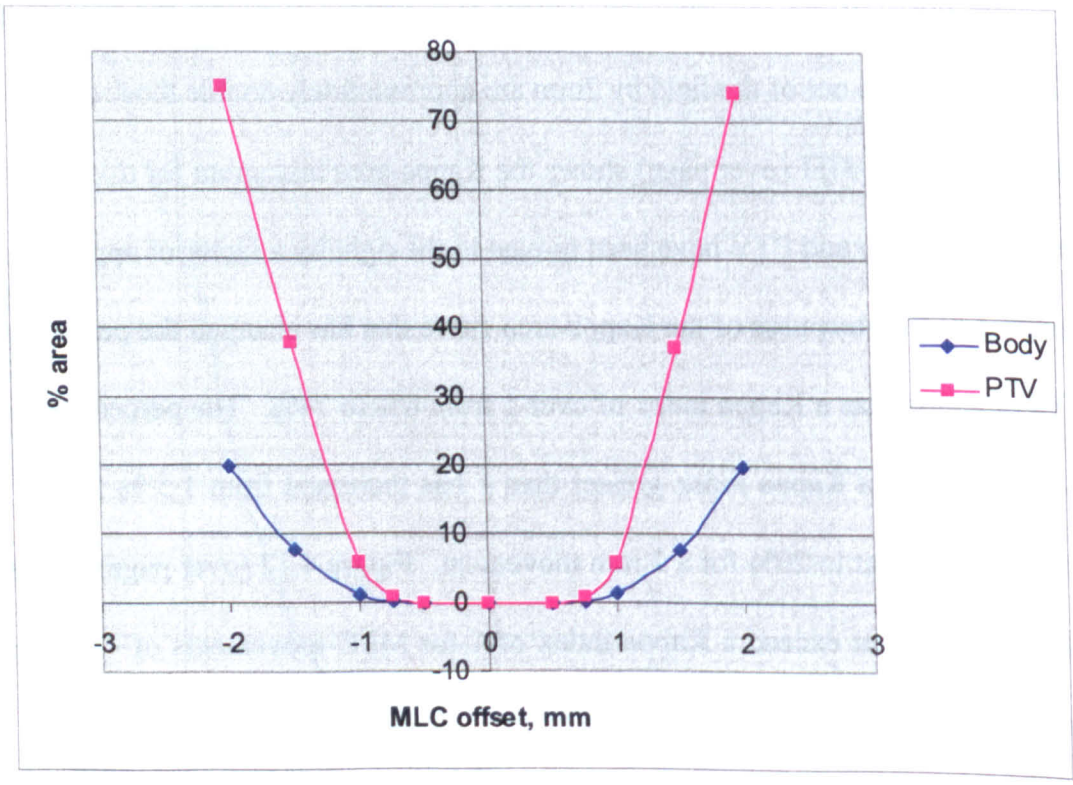



**Figure 4.12:** Percentage of the PTV and body areas in the isocentric slice that fail the Kappa 3%/3 mm test. Negative MLC offsets represent situations where both banks of MLC leaves are moved into the field, positive values indicate the MLC leaves moving out of the field. Failure is classed as any area with a Kappa index greater than 1.

shows that as the MLC movement is increased, the histogram lines move to the right by approximately the multiple of the MLC movement. For example the Kappa index at the 20% area point for the 'body' line move increases from 0.5 to 1 and then 1.5 as the MLC movement is increased from 0.5 to 1.5 mm. The 40% area point on the PTV line shows a similar trend, and for all lines the same basic shape can be seen, stretching out to the right as the MLC movement increases. Figure 4.14 (over page) shows the percentage area of the body and PTV that fails the Kappa test as the MLC movement is increased. The data in this graph are taken from the crossing points of the histograms in figure 4.13 and the $\kappa$=1 line. The rapid increase in failure area is due to the steep central portion of the histograms in figure 4.13. As the histogram curves move to the right, there is little change for movement of less than 0.5 mm followed by a rapid change as the central portion of the curves is encountered.

**Gravity-induced movement of MLC leaves**

Figure 4.15 (page 100) shows a dose difference plot for Patient H when the MLC leaves are allowed to move under gravity by 1 mm as the gantry rotates. The colour scaling is the same as for figure 4.8 showing the effects of both banks of MLCs having been moved, and the plot shows much less dose difference from the ideal distribution. Figures 4.16 (page 100) and 4.17 (page 101) show the Kappa-area histograms and failure-rate graph for various sizes of MLC gravity movement. Figures 4.18 (page 101), 4.19 and 4.20 (page 102) present the same dose difference plot, Kappa-area histograms and failure rate curve generated for Patient S's treatment.

**Segmentation**

The effects of segmentation algorithm are shown in figures 4.21 (page 103) and 4.22 (page 104), for Patient H's plan and Patient S's plan respectively. In each figure,

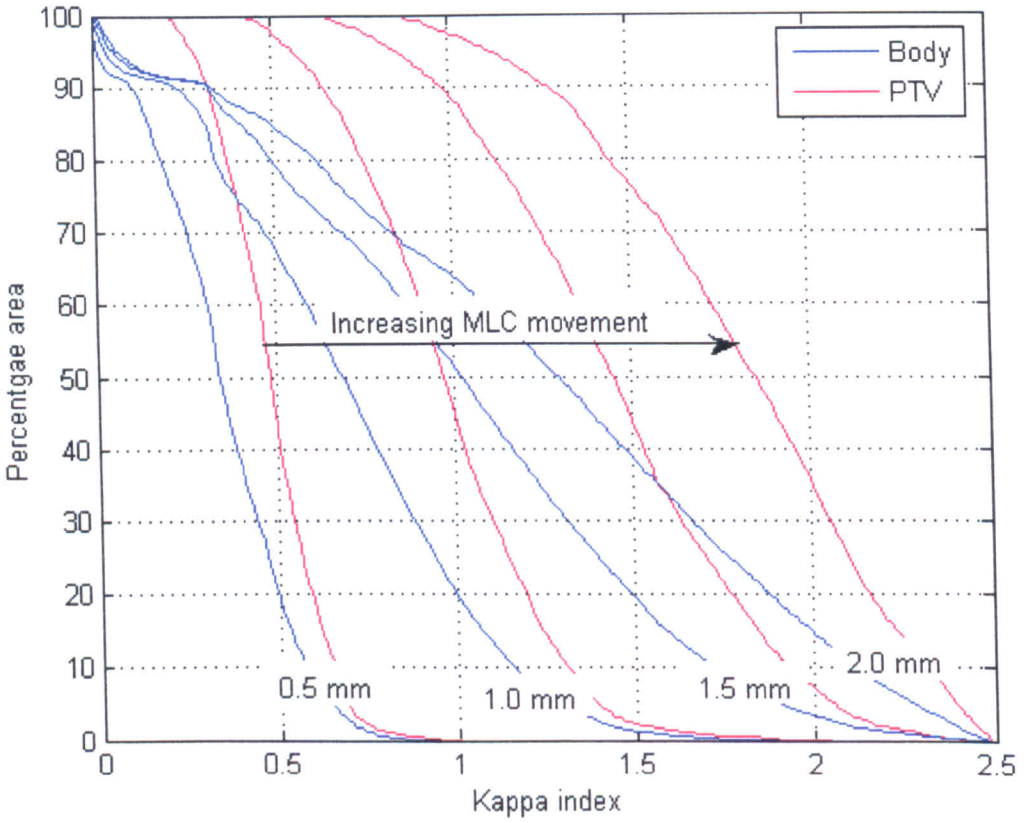



**Figure 4.13:** Kappa-area histograms for Patient S's treatment (head and neck plan) as the MLC offset is varied from 0.5 to 2 mm per leaf to make treatment fields larger.

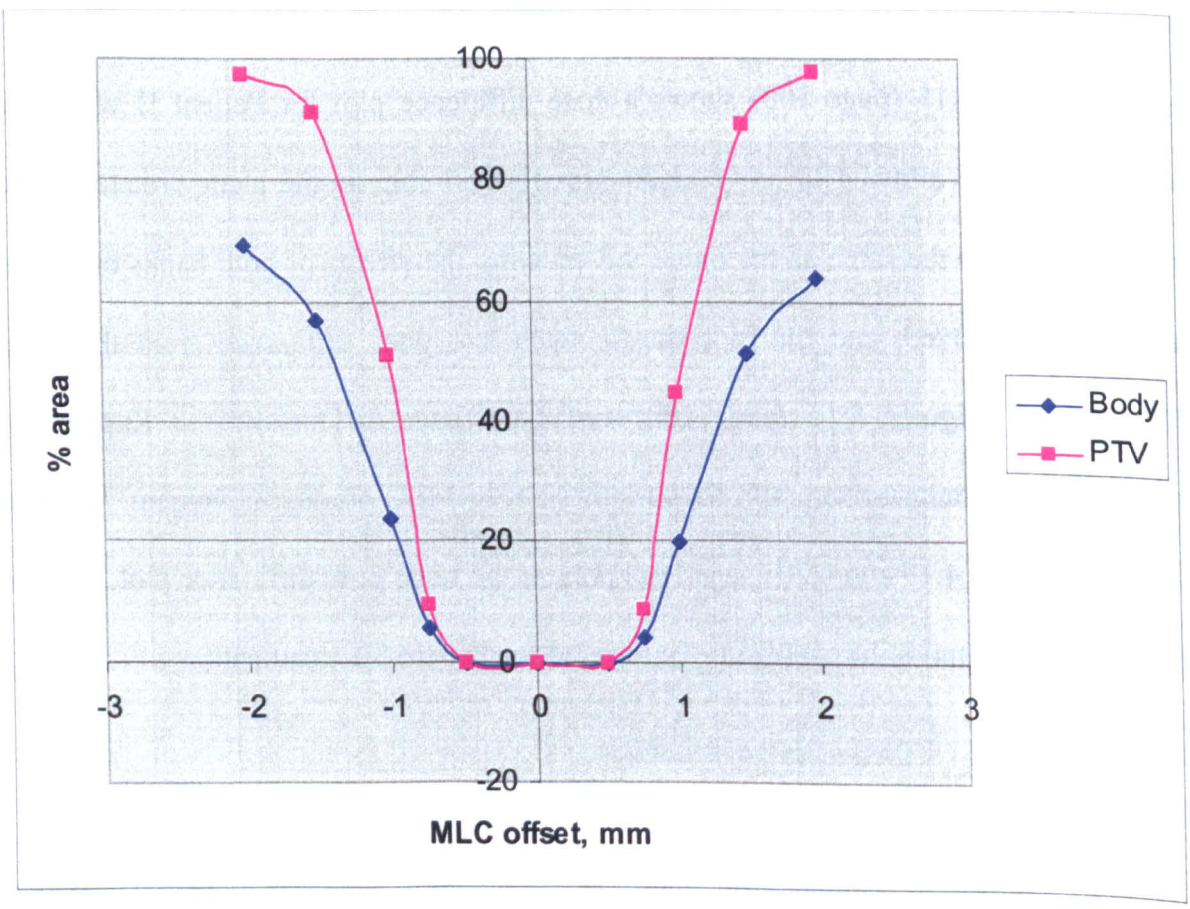



**Figure 4.14:** Graph showing the percentage areas of the PTV and body that fail the Kappa (3mm/3%) test for Patient S's (head and neck) treatment plan. The horizontal axis shows the offset of all MLC leaves, positive numbers representing the treatment fields becoming larger. The data points were derived from the crossing points of the curves in figure 4.13 with the Kappa=1 line.

the left hand column shows the dose differences, in percent of the isocentric dose, between a perfect delivery and one for which the MLC leaves are moved out of the field by 1 mm. The right hand columns show the effects of 1 mm of gravity motion. The colour scales within each figure are consistent, and the right hand columns show much less dose difference than the left, as noted above. The rows in figures 4.21 and 4.22 show the calculations performed on the same plan segmented in 5 different ways, namely: Slice, Standard, Standard Optimal, Platform and Platform Optimal. The figures show that there is very little difference in the dose differences created by plans segmented with these algorithms.

**Gravity-induced MLC leaf motion with an offset**

Figure 4.23 (page 105) shows for Patient H the dose difference against a perfect plan with a 2 mm gravity motion and 1 mm inward movement of all MLC leaves. The plot shows some hot and cold areas of dose, where the gravity motion is or is not counteracting the inward offset. Figure 4.24 (page 105) shows a Kappa analysis using 3% and 3 mm tolerances of the plot in 4.23, and figure 4.25 (page 106) shows the Kappa-area histogram analysis for varying levels of gravity motion, with half that motion subtracted as an offset. Figure 4.26 (page 106) shows the Kappa-area analysis for Patient S's plan under similar conditions.

**Dose output non-linearity**

Finally, figures 4.27 to 4.30 (pages 107 and 108) show the effects on a treatment plan of non-linear dose output. Figure 4.27 and 4.28 show the difference between a perfect machine and Patient H's plan delivered with Addenbrooke's machines LA7 and LA8 respectively, and figures 4.29 and 4.30 show Patient S's treatment delivered using machines LA7 and LA8 respectively. The plots clearly show an underdosing of most

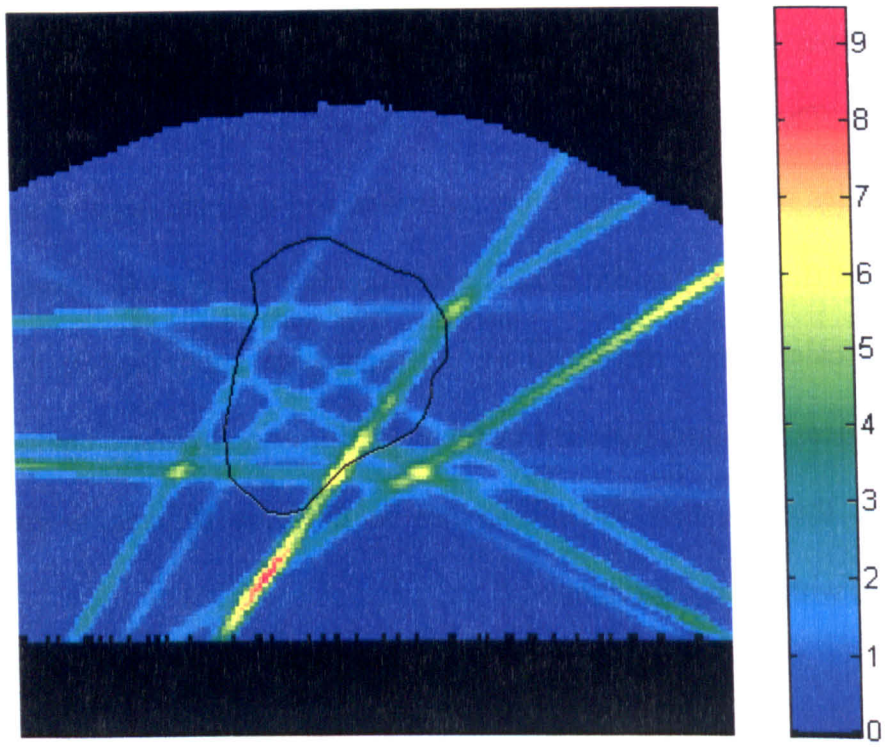



**Figure 4.15:** Dose difference plot for Patient H showing the difference in the dose distribution created by a 1 mm gravity motion of the MLC leaves. The colour scale is in percentage of the isocentre dose and the PTV structure is shown as a black outline.

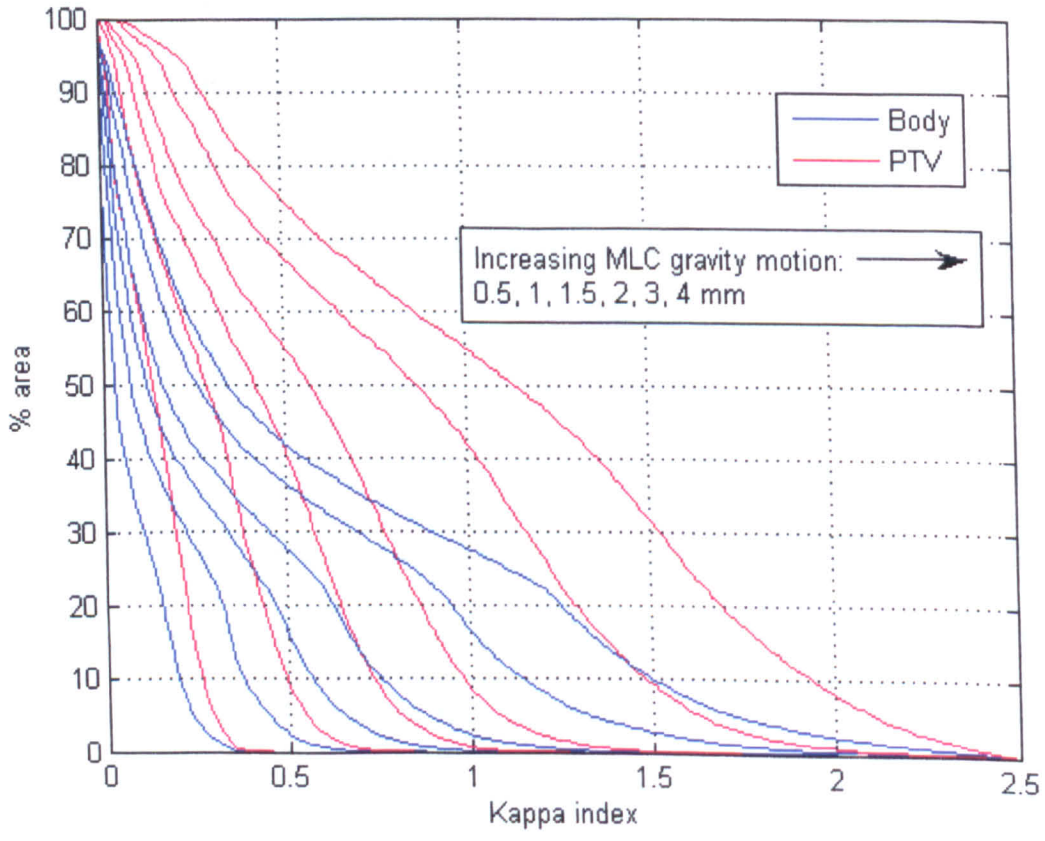



**Figure 4.16:** Kappa-area histogram showing the dose differences for various gravity movements of the MLC leaves for simulations of Patient H's treatment plan. From left to right the curves show 0.5, 1, 1.5, 2, 3 and 4 mm movements.

areas when LA7 is used and an overdosing if LA8 is used, however the effect is small, always under 0.4% for plan H and 0.8% for plan S.

## Discussion

### MLC leaf positioning

Figure 4.8 shows an example of the specific dosimetric errors that are introduced into a dose distribution by systematic changes in MLC leaf positioning. As each MLC leaf moves in or out of the field, it is moving the location of the field's penumbra in that location, which is very steep at the leaf-end itself. This causes narrow rods of over or under-dosing that penetrate the entire 3-dimensional dose distribution. Where these rods overlap, there is potential for significant increase or decreases in dose. The rods of dose difference occur at every point throughout each beam's-eye-view projection that an MLC leaf-end is located. The gradient of the beam-edge penumbra (as shown in figure 3.1) is approximately 12% per millimetre at the 50% level. Thus a movement of the MLC leaf by 1 mm causes a rod of overdose with a maximum intensity of +12%. This percentage is of the intensity of that particular beamlet. IMRT uses many low-dose beamlets, so as a proportion of the total treatment each rod of overdose is small. For Patient H's treatment, using 50 beamlets, each rod of overdose has an approximate contribution of 12/50=0.24% of the total dose distribution (assuming that all segments are of approximately equal dose). At first it would seem that these rods of 0.24% overdose would not cause a problem to the overall treatment. However, where these rods overlap, the dose increase is built up, causing the 9-10% hotspots shown in figure 4.8. The dose difference plot shown in figure 4.8 shows that not only do these rods from different beam angles combine but that rods from the same direction also combine to form larger rods of higher dose difference. From any particular beam angle there are rods of overdose that are less than 1%, which would equate to single overdose rods, and

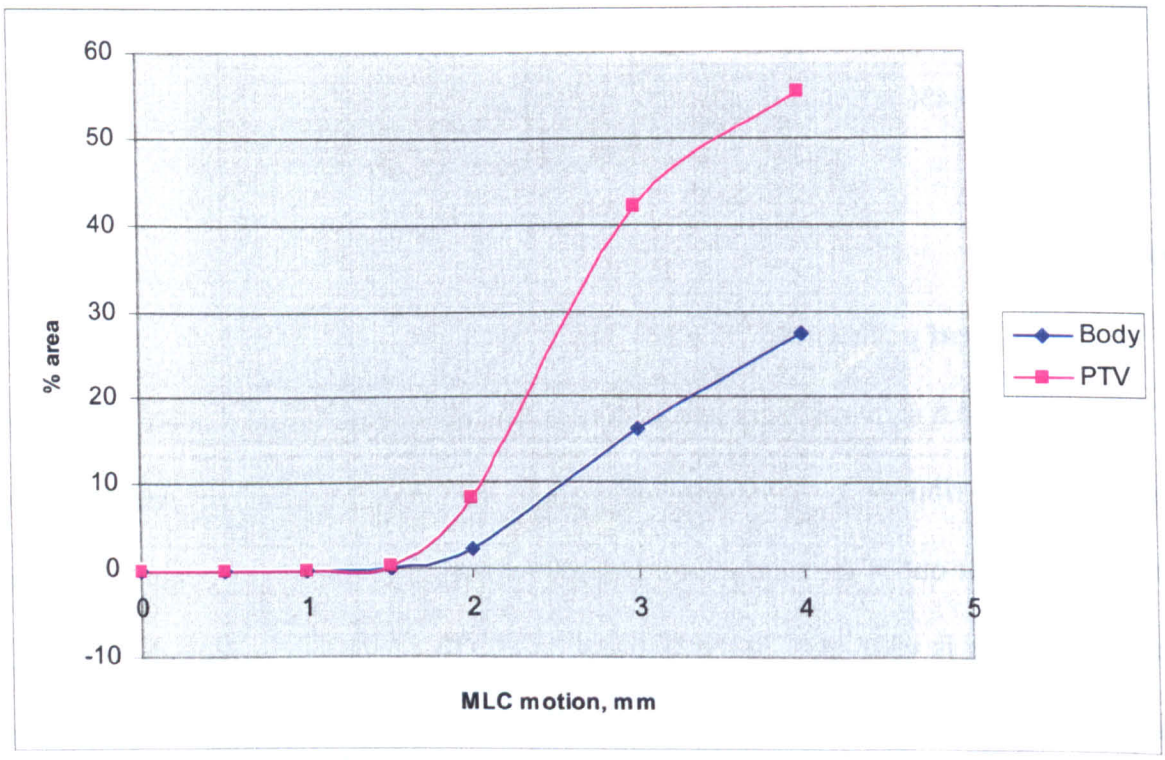



**Figure 4.17:** Graph showing the percentage of the body and PTV structures that have a Kappa index greater than 1 as the gravity motion of the MLC leaves is increased. The data for this graph are taken directly from the Kappa=1 point in figure 4.15.

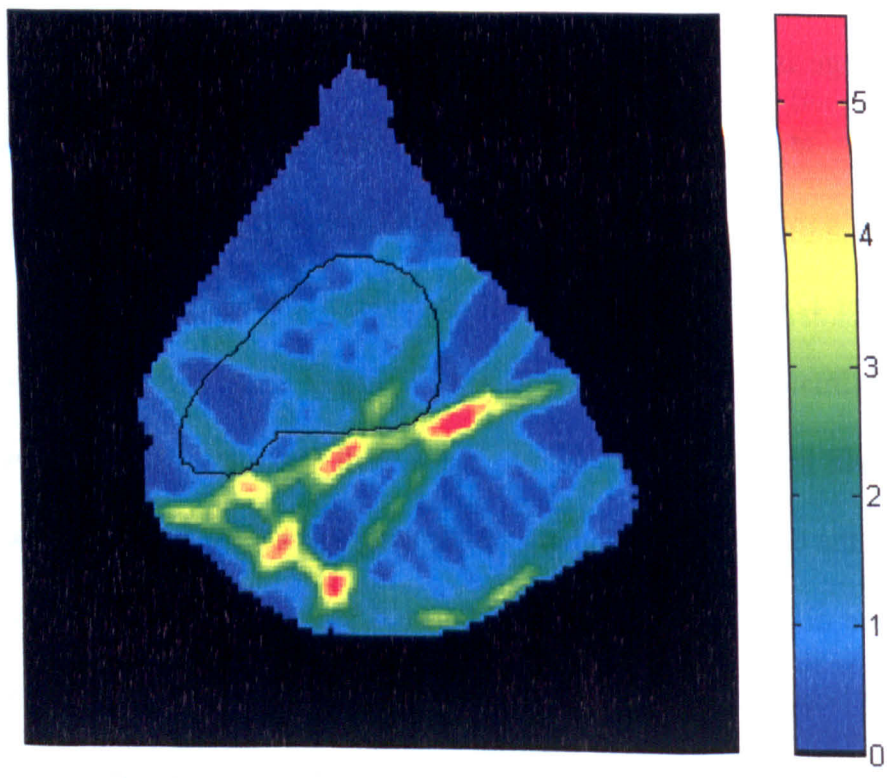



**Figure 4.18:** Dose difference plot for Patient S's plan, generated by a 1 mm gravity motion of the MLC leaves. The colour scale is in percentage of the isocentre dose and the PTV is marked as a black outline.

also rods of 2-3% and even 4-5% overdose. It is when these larger rods combine that the overall 9-10% hotspots are caused. Many of the 2-5% overdose rods fall along or near to the edges of the target volume. Why this should be so is easy to see if one considers the likely fluence profile for a particular beam direction. In general, MLC leaf ends need to be placed wherever there is a step in the fluence map for a beam direction. In building up a modulated fluence profile, it is impossible to have a step in the fluence without having an MLC leaf positioned there to shield the lower part of the profile on at least one segment. Thus it is clear that the largest steps in a fluence profile will have, when summed over a set of segments that create a beam, the largest number of MLC leaf-ends positioned there, and these areas will experience the summation of several beamlet edges errors. Outside of the beam edge (as defined by the summation of the beamlets), the fluence is zero. It is unlikely that the fluence at the edges of the profile will be low; the edges of a fluence map are in fact likely to have high fluence in order to boost the coverage of dose over the edges of the target. Thus at the edges of a fluence map we see the largest step in fluence, thus the largest number of MLC leaf-ends and consequently large errors in dose if there is a systematic leaf positioning error.

The question that arises, if the largest dosimetric errors are on beam edges and therefore in high dose-gradient areas, is whether 9-10% hotspots are actually a 'problem' in terms of radiotherapy treatments and planning. Distance-to-agreement and treatment planning margins share the concept of dose coverage to target volumes and dose sparing to organ at risk structures grown to a planning risk volume. Margins allow for a dose distribution to move without affecting the dose delivered to structures outside of tolerable levels, and thus the analysis of distance-to-agreement as an acceptance criterion depends on margins having been added to cover whatever inaccuracy has caused two dose distributions (real and planned) to differ. Margins are added at the boundaries of structures where there will inevitably be steep dose gradients. Figure 4.9

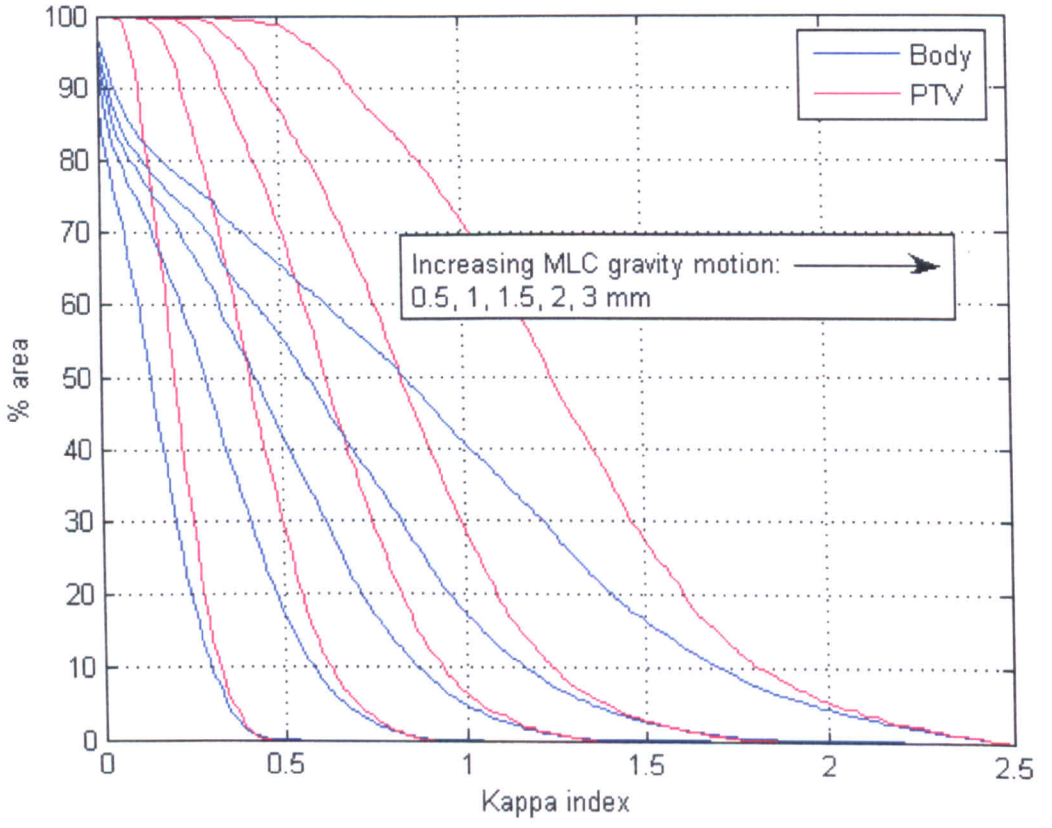



**Figure 4.19:** Kappa-area histogram curves for Patient S's treatment plan as a function of gravity motion of the MLC leaves.

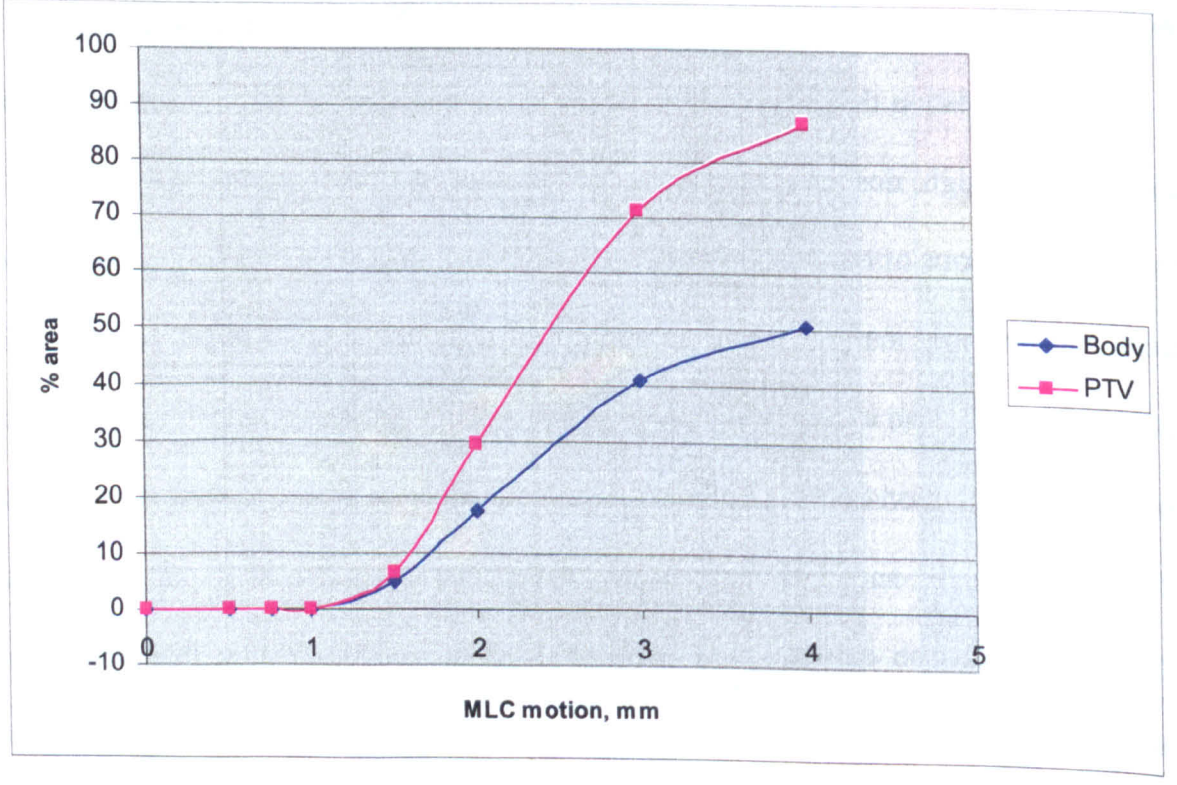



**Figure 4.20:** Graph showing the increase in percentage of the body and PTV structures that have a Kappa index of more than 1. The points on this graph correspond directly to the crossing point of the Kappa=1 line on the graph above.

shows a Kappa index analysis of the dose distribution shown in figure 4.8. Many of the largest and highest hotspots from figure 4.8 show a low Kappa index precisely because they are in high dose-gradient areas and the DTA is therefore small. Areas along the edges of beams will therefore tend to have lower Kappa index values and may pass whatever tolerances are set for testing a dose distribution. The tolerance levels chosen for the Kappa analysis shown were 3% and 3 mm. These are smaller values than are typically used for IMRT verification as a whole, with Addenbrooke's using 4% and 4 mm for its IMRT program. However, the 4/4 tolerance levels exist as an overall set of values for validating all of the potential errors associated with IMRT. The topic of interest here is purely those errors that are introduced by systematic errors in the behaviour of the treatment machine. The tolerances used for overall IMRT verification also have to include such issues as the calculation accuracy of the treatment planning system and accuracy of the equipment used to verify a treatment plan. The issues under investigation here compare one plan directly to another, and so in short, less is being evaluated and therefore smaller tolerances are appropriate. Conversion to other tolerances, assuming that the same numeric value is chosen for the DTA and dose difference (see chapter 5), is relatively easy, with values requiring a simple multiplier.

Figure 4.9, then, shows that many of the hotspots from figure 4.8 give only small values of Kappa. However, dose differences inside the target volume, where the dose distribution is intended to have a low gradient, have large DTA values and therefore the overriding value that shows up in the Kappa index is that of the dose difference. In this case, dose differences of 4-5% give Kappa indices of 1.3 to 1.7. The maximum value of Kappa in figure 4.9 is 1.9, which comes from a dose difference of 5.7%. Because that is inside the target volume, the DTA is large and the Kappa index is subsequently large, indicating a failure of the tolerance levels. In fact, figure 4.9 shows that most of the areas that fail the Kappa test are inside the target volume, but figure 4.8

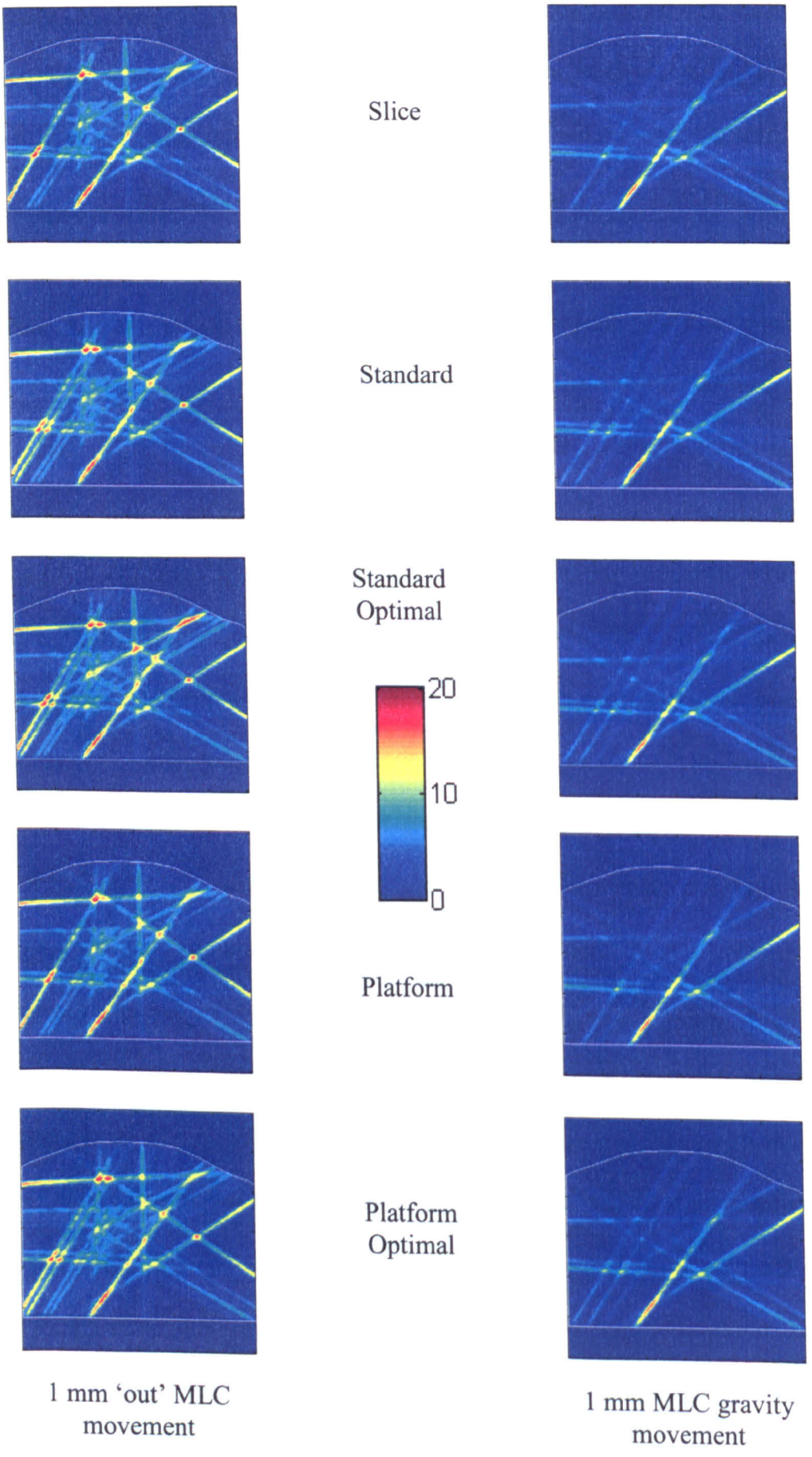



**Figure 4.21:** Dose-difference plots for Patient H's treatment plan as a function of segmentation algorithm. The left-hand column shows MLC offsets of 1 mm out of the field, the right-hand column shows gravity movements of 1 mm. The 5 rows show the 5 different segmentation algorithms available to ImFAST, as labelled. The colour scale is percentage of the isocentric dose and is the same for all images.

indicates that many of the largest and worst hotspots are outside of the volume. This clearly shows that analysis of a dose distribution depends critically on the tool used.

Figure 4.10 demonstrates that the area of the target structure that fails the Kappa index test is approximately 6%. A change in the Kappa tolerance levels as noted above would stretch the Kappa-area histogram to the left or the right. A tolerance of 4% and 4 mm, for example, would move the curve such that the current value of $\kappa$=1.33 would become the new $\kappa$ =1. The area failing the Kappa index would then be just 1.5%. However, if a tighter tolerance is used, such as 2% and 2 mm, the current $\kappa$ =0.66 is of interest, in which case nearly 40% of the target structure fails the Kappa index.

Figure 4.11 shows that when the MLC offset is increased to 2 mm per leaf, the Kappa-area histogram is stretched to the right. The whole curve is moved to approximately double the previous Kappa index values – for the PTV, the 75% level moves from a Kappa index of 1 to almost 2, and the 25% level moves from around 0.75 to nearly 1.5. This is not an unexpected result, as the shapes of a beam penumbra is approximately straight in its middle section, and thus moving the penumbra 2 mm will produce about double the dose difference as a 1 mm move. In areas of the dose distribution where the dose gradient is approximately constant, a doubling of dose difference will double the distance to agreement and the Kappa index will approximately double everywhere. The exception is the tail of the Kappa-area histogram, which does not move to double the value between figures 4.10 and 4.11. This is due to the composite nature of the Kappa index – in some areas where previously the DTA gave a lower Kappa index, the dose difference may now do so, and vice versa. The non-uniform gradient nature of dose distributions means that the tails of Kappa-area histograms will not always increase in proportion to the dose differences.

Figure 4.12 shows how the percentage area that fails the Kappa=1 test changes as a function of MLC offset, and shows that the effect is not linear. Indeed, for MLC

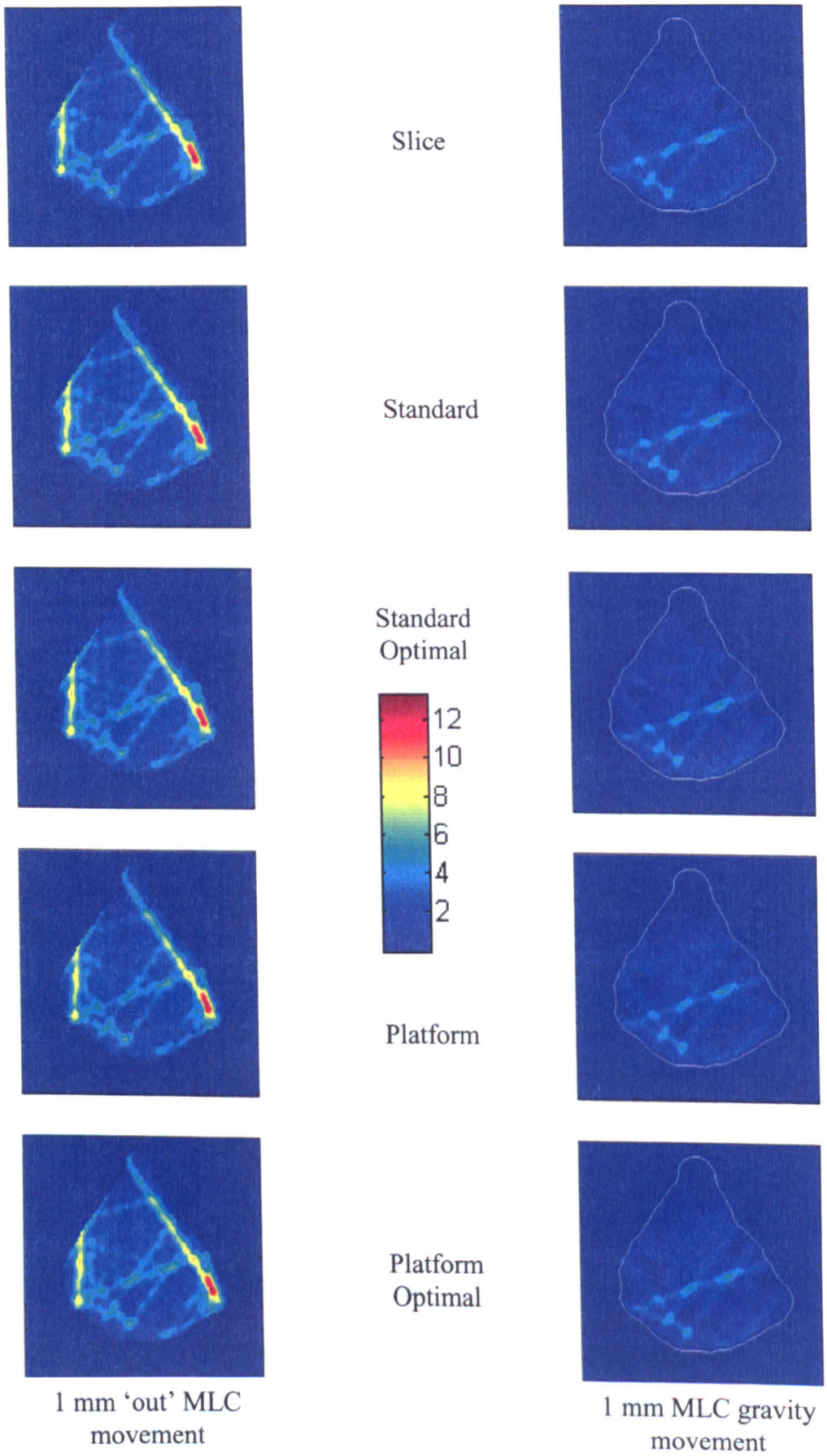



**Figure 4.22:** Dose difference plots for Patient S's treatment plan as a function of segmentation algorithm. The left-hand column shows MLC offsets of 1 mm per bank out of the field, the right-hand column shows 1 mm gravity movement. The 5 rows show the 5 algorithms used, as labelled. The colour scale is a percentage of the isocentre dose and is the same for all images.

offsets less than 1 mm, the percentage of the target volume that fails the Kappa index test is below 10%. While it would be ideal for none of the area to fail, levels as low as 10% may be acceptable to oncologists. However, for MLC offsets larger than 1 mm, the percentage area that fails rapidly increases. This is due to the steepness of the Kappa-area histogram through its central portion. As the MLC offset increases, the curve is stretched to the right. Not only do small changes in the horizontal position of the Kappa-area curve make large differences to the percentage failure, but the curve is stretched to the right by increasing amounts as the MLC offset gets larger, due to the multiplication stretch of the curve rather than a lateral movement. Hence figure 4.12 shows very steep increases in the failure rate of the dose distribution. The PTV structure shows a much faster increase in failure for the reasons discussed above – much of the high Kappa index values are contained within the low dose-gradient areas. The curve showing the 'body' failure rises at a lower rate because it encompasses a larger volume with little more failure than that within the PTV itself.

The analysis of dose differences, Kappa index plots and failure rates have so far focussed on a single treatment plan. Figures 4.13 and 4.14 show the Kappa-area histograms and Kappa failure rates for the second patient, Patient S. The treatment plan for Patient S was more complex than for Patient H, with 89 segments rather than 50, and treatment to a more complex target in the head and neck rather than the relatively simple target shape and location in the abdomen. As such, the effects of MLC offset are more drastic than for Patient H's treatment. For Patient H, the failure rate of the PTV structure was around 6% for a 1 mm MLC offset. For Patient S, that rate increases to over 40% of the PTV falling outside of $\kappa = 1$. To achieve a similar failure rate of 6%, an MLC offset of just 0.6 mm is allowed. The more rapid increase in failure rate is due to the more complex geometry of the treatment beams and of the patient contours; for Patient H, the dose differences could be seen to decrease as the beams penetrated deeper

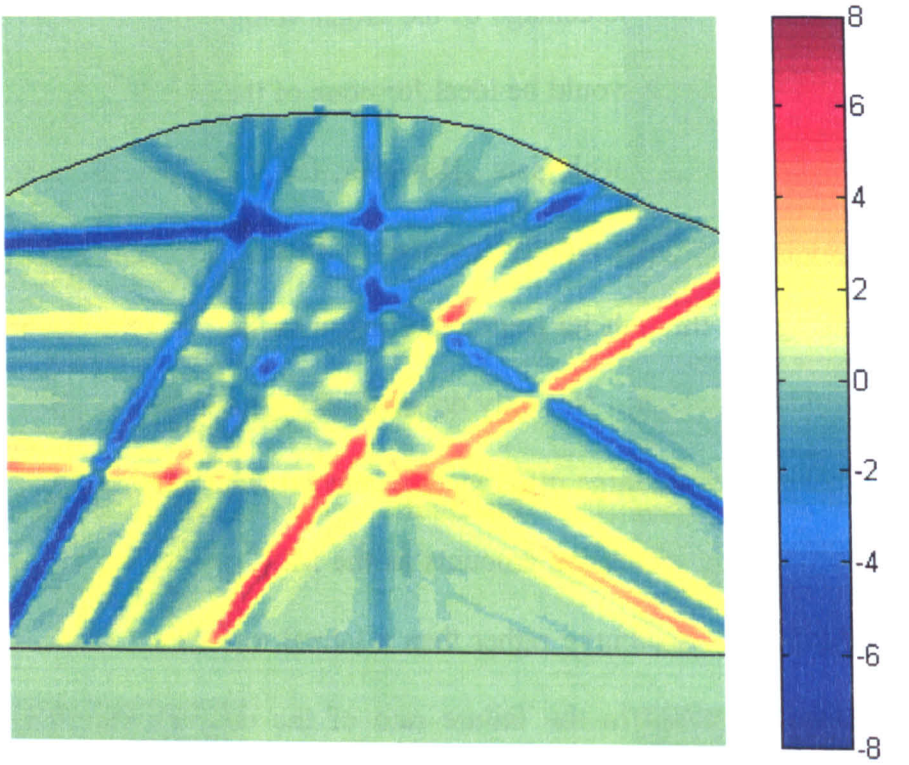



**Figure 4.23:** Dose difference plot for Patient H's treatment plan. The MLC leaves experience 2 mm of gravity motion, but are calibrated to be 1 mm into the field. The colour scale is dose difference as a percentage of the isocentre dose.

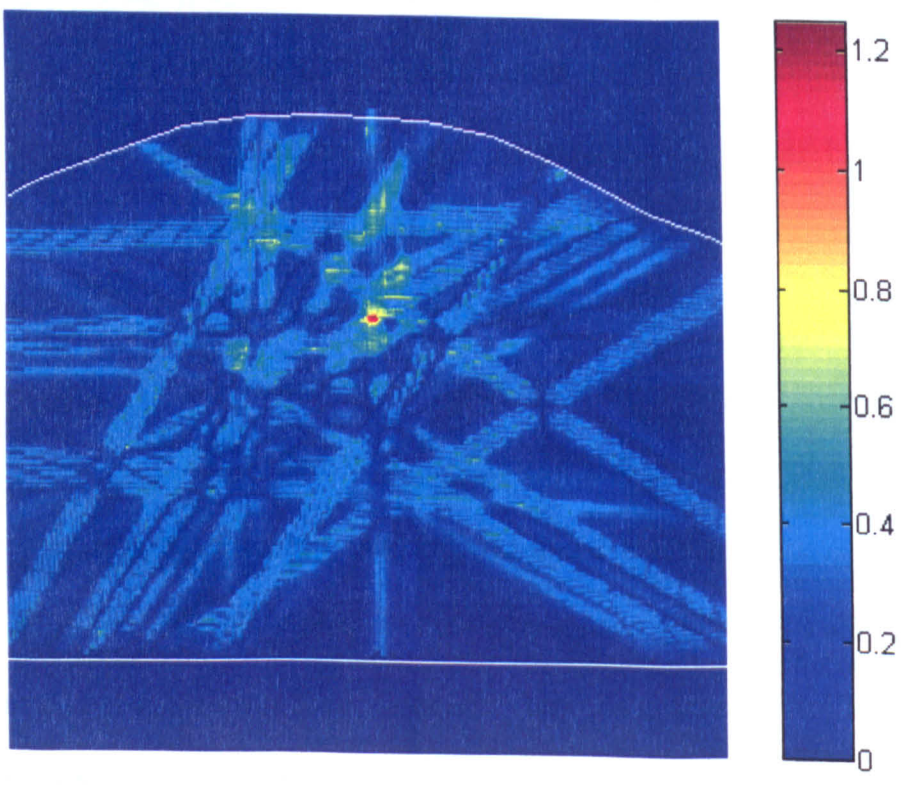



**Figure 4.24:** Kappa index analysis of the above dose difference plot, using 3% and 3 mm tolerance levels.

into the patient. For Patient S, there are no areas of the patient deeper than around 10 cm, with the isocentric axial slice being in the neck. Thus the dose differences are greater and add up more quickly than do those for Patient H.

It is clear then that the dose differences and analysis of treatment accuracy depends critically upon the specifics of the treatment site and the treatment fields themselves. Perhaps the key issue is the degree of modulation involved in a treatment plan. A head-and-neck case involves small volumes which are not very deep, with rapidly changing patient contours and the inclusion of airways and bony structures. A prostate plan involves deeper areas within nearly homogenous areas of tissue, and will therefore require a smaller degree of modulation to achieve an acceptable dose distribution. This results in a smaller number of segments, as seen with Patient H and Patient S, and these segments will cover a larger area. Thus each segment is larger, and a systematic change in the positioning of MLCs will have a smaller effect on these larger segments. The head-and-neck plan is more sensitive to MLC positioning because it has more and narrower segments.

**Gravity-induced MLC movement**

Figures 4.15 to 4.20 show the effects of gravity-induced motion of the MLCs on the dose distributions. All of the dose differences are positive (overdose) because of the 'swung-in' calibration of the MLC leaves at gantry 0 and the subsequent opening of the leaves as the gantry rotates. The dose difference plots (4.15 and 4.18) show much smaller differences than for the same motion applied to both MLC banks because the gravity motion affects only one bank of leaves unless the gantry is very close to 180 degrees. The Kappa-area histograms and the failure-rate graphs show that an approximate doubling of the distance moved is required for a similar change in the dose distribution. This shows that the effect of one bank of leaves opening by 2 mm is

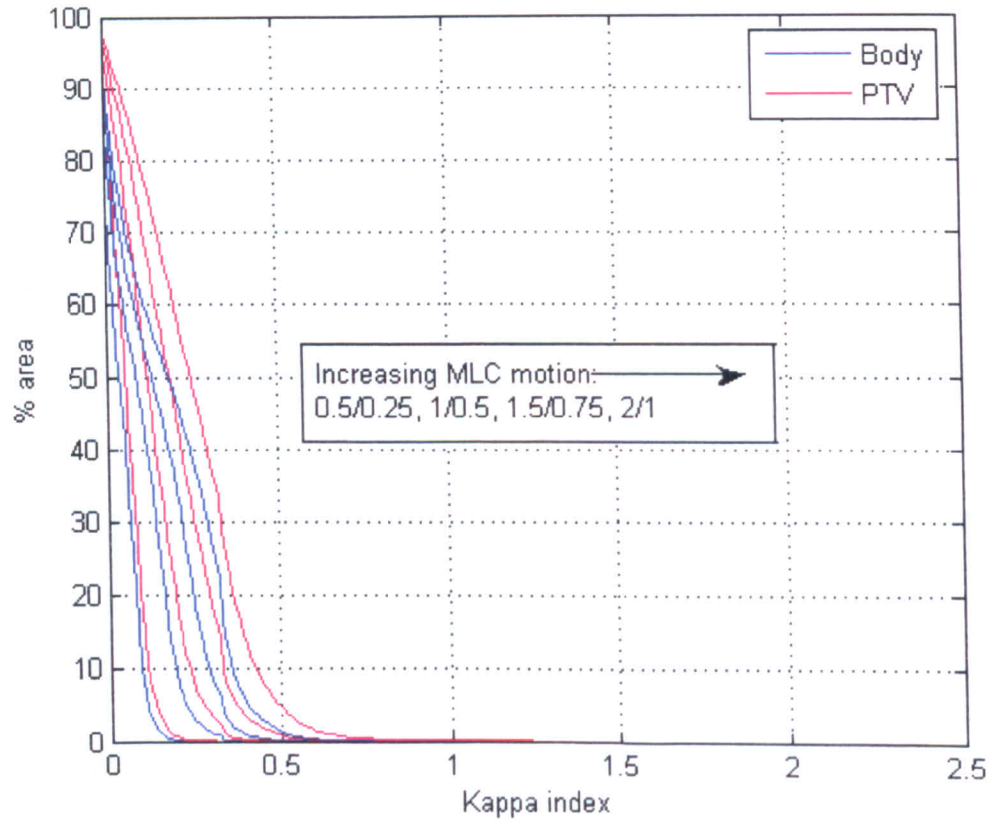



**Figure 4.25:** Kappa-area histograms for Patient H's treatment plan for situations in which the MLC leaves are deliberately calibrated to be too far into the field by half of the gravity motion quantity. The four lines show gravity motions of 0.5, 1, 1.5 and 2 mm with offsets of 0.25, 0.5, 0.75 and 1 mm respectively.

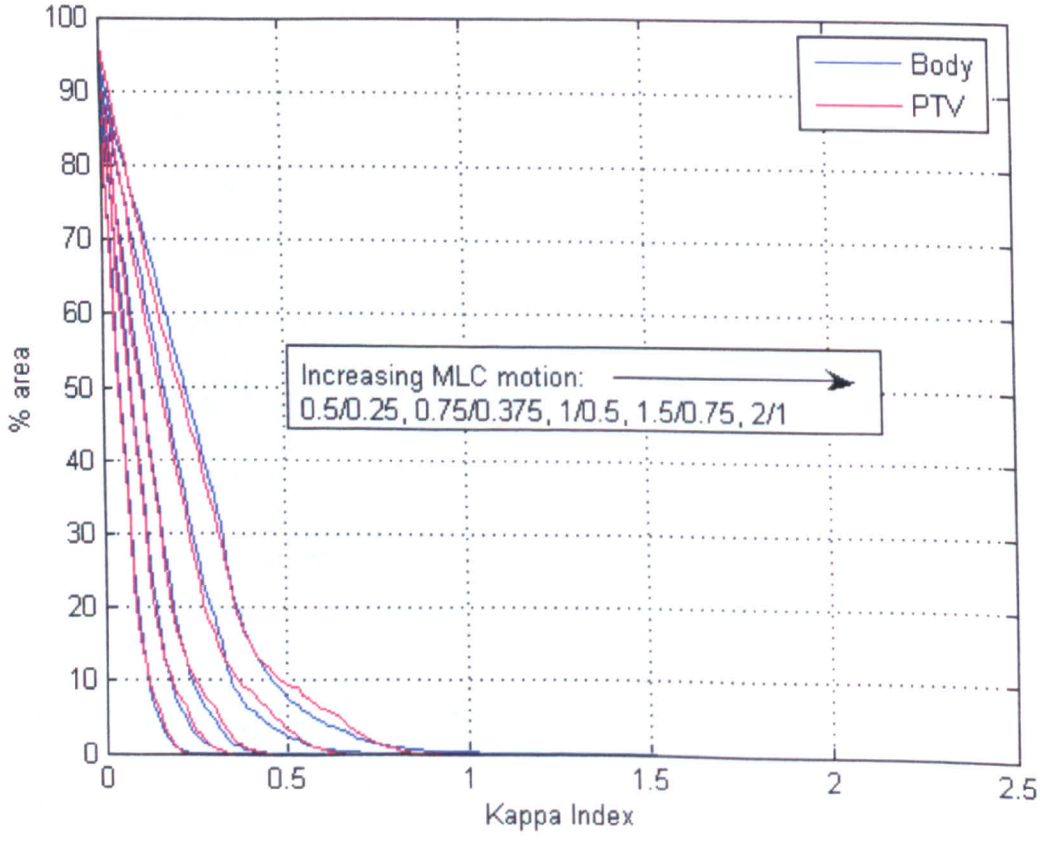



**Figure 4.26:** Kappa-area histograms for Patient S's treatment plan, showing the distribution of Kappa indices as the gravity motion of the MLC leaves is increased from 0.5 to 2 mm with deliberate MLC offsets of 0.25 to 1 mm respectively.

equivalent to both banks opening by 1 mm. The overall field size and shape will be the same, with one field offset by 1 mm compared to the other. Such overall movements of the field therefore do not have much of an effect on the dose distribution, and it is the increase in size that creates the dose differences. The graphs of failure rates show that the effects of gravity motion on the MLC leaves do not start to create Kappa index failures (for 3% and 3 mm tolerance) until the movement reaches around 1.2 mm, with a 1.5 mm movement causing nearly 10% failure for Patient S. As before, the failure rate for a tolerance of 2% and 2 mm can be read from the Kappa-area histograms by looking at the $\kappa$ =0.66 line, indicating that a 1 mm movement of the leaves would lead to failure of over 5% for Patient S. These figures therefore show that an MLC positioning tolerance (in terms of gravity-based movement) of half the Kappa distance tolerance is required if failure rates are to be kept below 10%.

**Segmentation**

The dose difference plots presented in figures 4.21 and 4.22 show that there is very little difference in the effects of MLC movement as the segmentation algorithm, and therefore the delivery of segments, is altered. Interestingly, the algorithm that shows the worst hotspots is one of the so-called 'optimal' ones, Standard Optimal. This is an implementation of the sliding window technique. Although there is only one way to segment a 1D profile using sliding window (figure 1.10), mechanical constraints with the MLCs mean that there are many ways of combining the 1D windows into deliverable 2D apertures. The Optimal algorithm generates many hundreds of combinations and selects the best based on minimum treatment time.

As mentioned previously, dose differences can only occur where an MLC leaf end is positioned and any step in a fluence map will require an MLC leaf end to be positioned there at some point in the treatment delivery. Thus is it not unexpected that

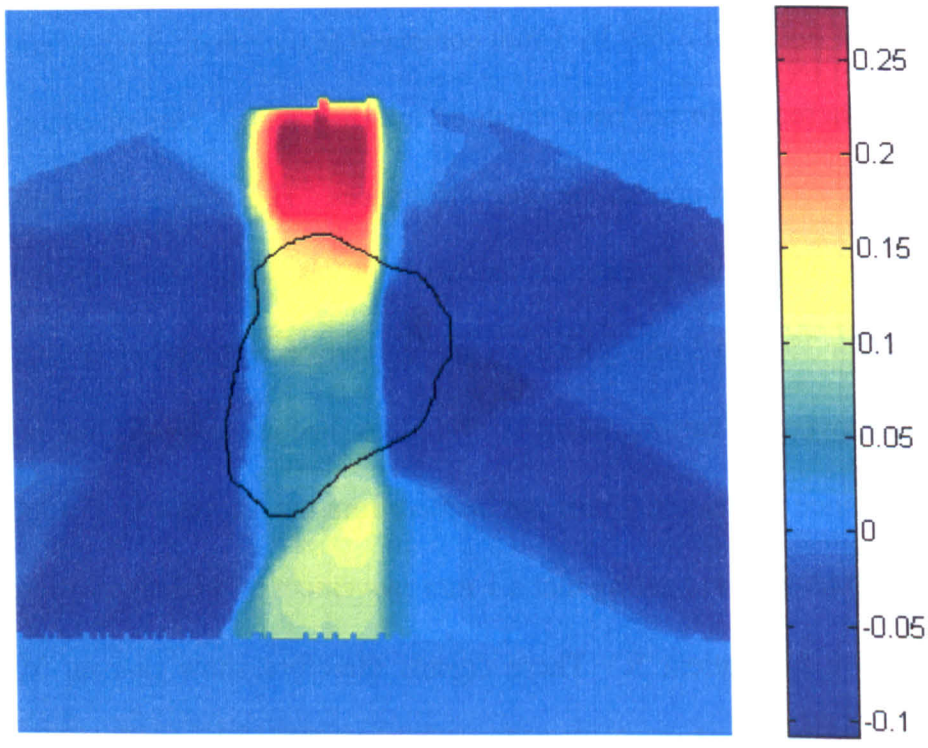



**Figure 4.27:** Dose difference for Patient H's plan between treatment on a linac with perfectly linear dose output and treatment on Addenbrooke's machine LA7. The colour scale is in percentage of the isocentre dose.

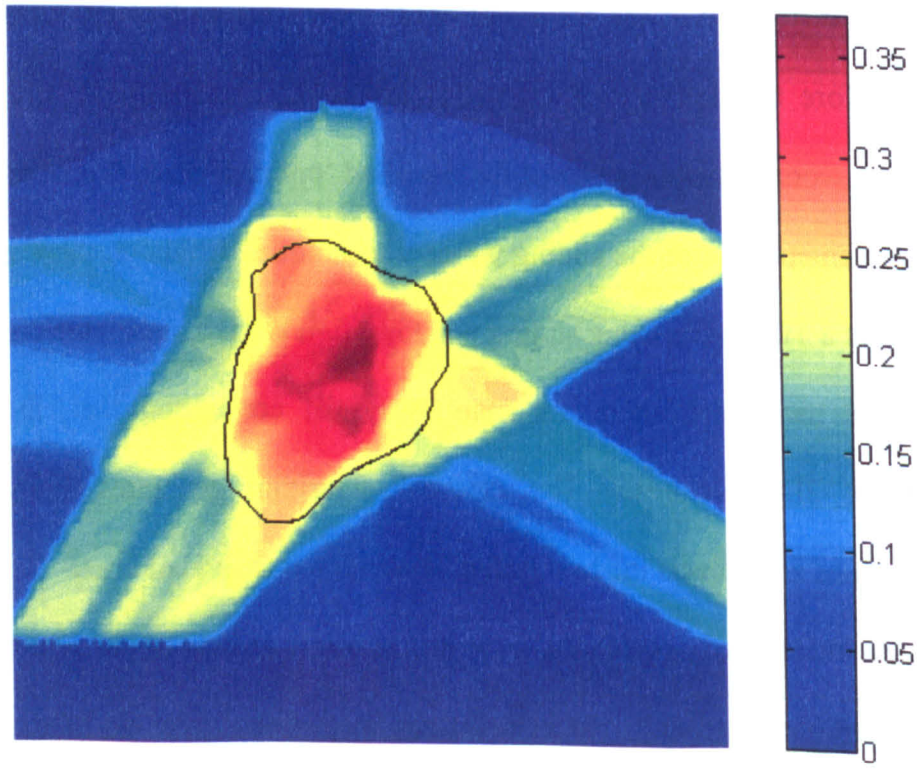



**Figure 4.28:** Dose difference as for figure 4.27, but using Addenbrooke's machine LA8 for treatment.

different algorithms therefore produce the remarkably similar plots shown in figures 4.21 and 4.22. The subtle differences are caused by the algorithms sometimes having to divide segments into two because of the mechanical constraints and therefore positioning leaf-ends where there is no step in the fluence profile. This splitting of fields is dependent upon the exact configuration of a fluence map and the ways of segmenting it according to each algorithm and is not easy to quantify. However, for the plans chosen here the differences between the algorithms, in terms of the different dose errors, is small compared to the overall errors introduced by the necessity of MLC placement in the first place. It may be possible to devise an all-round optimal segmentation algorithm, although this work shows that any benefits will likely be small.

**Gravity-induced MLC leaf motion with an offset**

After the gravity motion effect on the MLC leaves was discovered, it was hypothesized that the effects of gravity motion on the MLC leaves could be partially compensated for by a deliberate miscalibration of the MLC leaves. Since the MLC leaves are always calibrated at a gantry angle of 0 and the field sizes will always get larger with gantry rotation, a small offset into the field could give fields that are slightly too small at gantry angles of 0 and slightly too large at other gantry angles. Figures 4.25 and 4.26 show Kappa-area histograms for the case of various gravity motions and deliberate miscalibration of half that motion. Comparing back to the Kappa histograms of figure 4.16 and figure 4.19 for gravity motion only, it is clear that the curves have been shifted to the left, and that for gravity motion as high as 2 mm, the failure rate of a Kappa 3%/3 mm analysis is less than 1% of the body or PTV for Patient S and even smaller for Patient H's more simple plan.

All of the above data shows some interesting effects. Although the effects of gravity motion on a plan are less pronounced than those for MLC offsets, the gravity

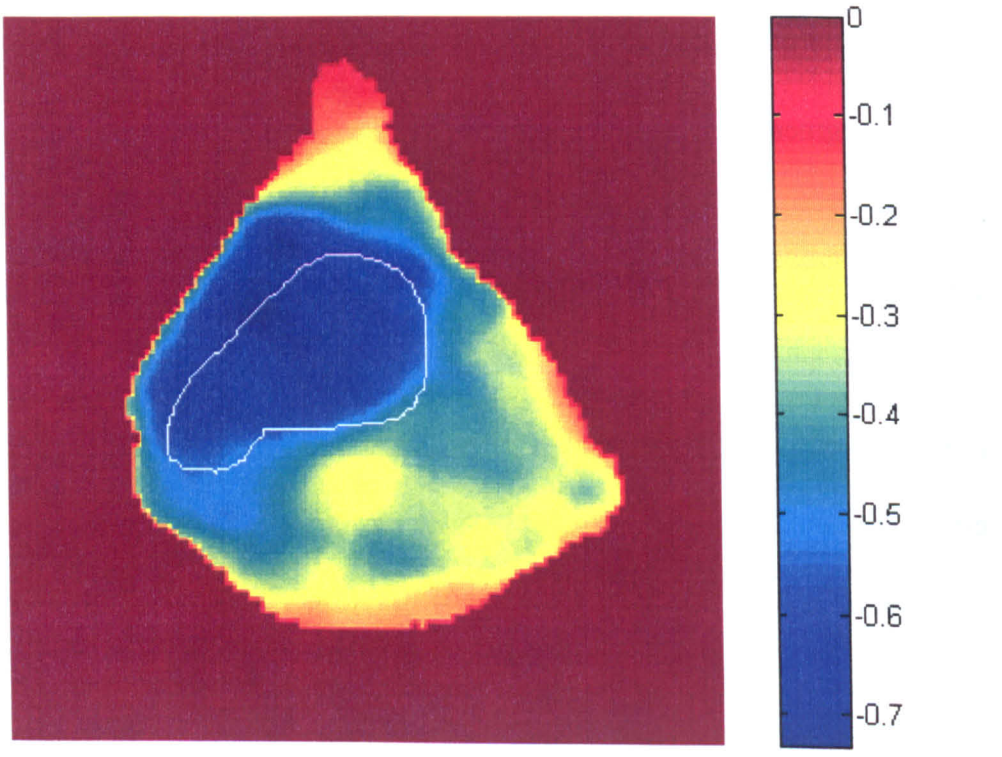



**Figure 4.29:** Dose difference for Patient S's treatment when treated used Addenbrooke's machine LA7 rather than a machine with perfect dose linearity. The colour scale is in percent of the isocentre dose.

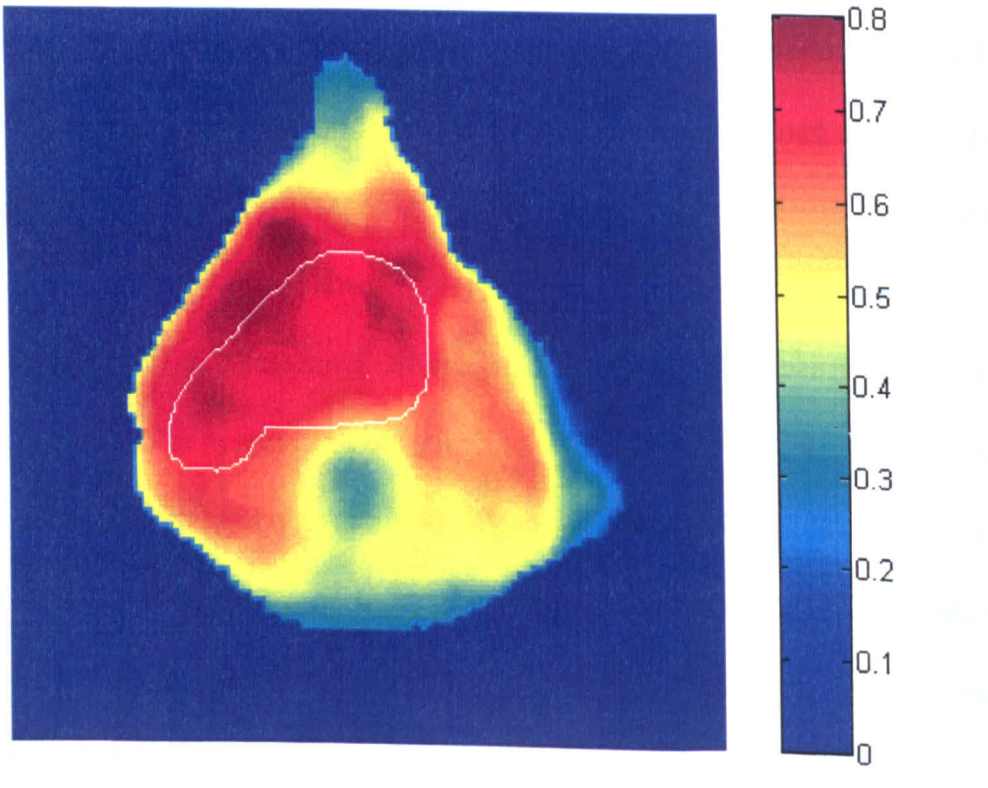



**Figure 4.30:** Dose differences for Patient S's treatment when treated using Addenbrooke's machine LA8 rather than a machine with perfect output linearity. Colour scale as for figure 4.29.

motion will be present on every plan treated on a particular linac. The MLC offset may change as the MLC leaves are calibrated or reset, and could vary from -2 to to +2 mm. The probability of having such large offsets on both banks of MLCs is small if an effective QA procedure is in place. However, if the techniques used for calibration of MLCs are poor, such as relying on alignment to a light field, then miscalibration of this order is not inconceivable. The MLC measurement software written in order to identify the MLC motions can be used as an MLC measurement tool for regular calibration and checking of the MLC leaves. The use of film and simple software provides an easily traceable QA route that is accurate and fast. Gravity-induced motion of the MLC leaves cannot be removed, and will be present for all treatments. However, the effects could be minimized by using varying collimator angles for different beams in a treatment plan. Treatment planning systems do not allow a mixture of collimator angles for the segments that make up one beam, as the optimization of the fluence map would be more complex and is not supported. However, different collimator angles can be used for different beam angles, and angles near collimator 0 can be avoided to minimize the component of gravity that acts parallel to the MLC direction of travel. Additionally, this work shows that a deliberate miscalibration of the MLC leaves can significantly reduce the effects of gravity motion. Again, the MLC measurement software developed here can be useful in this process, as the measurement precision of the system is around 0.17 mm, which would enable gravity motions as small as 0.34 mm to be offset.

**Dose output non-linearity**

The effects of non-linear dose output are obvious – if 10 monitor units is not the same as 2 times 5 monitor units then there will be a dosimetric difference between different treatment plans that use different number of monitor units per segment. Figure 4.7 shows that the linacs at Addenbrooke's have linearity within 2.5% with most

numbers of monitor units giving linearity better than 1%. Figures 4.27 to 4.30 show that the maximum dose error for Patient H is nearly 0.4%, while that for Patient S is 0.8%. For Patient H the effect of some beams underdosing and others overdosing is clear in figure 4.27 (using machine LA7). While these beams largely cancel in the target volume to give very small errors, for the other three plots all of the beams are either hot or cold. There are two reasons for this. Firstly, figures 4.7a/b show that the dose linearity curves are slowly changing such that they do not flip from above to below 1 very much. Thus there are whole ranges of monitor units for which the dose is either high or low, rather than the difference being randomly distributed. The second reason is the segmentation algorithm. As described in chapter 1 and shown in figure 1.9, the segmentation algorithms used in Xio and Corvus divide the fluence into equal steps, which leads to segments of equal numbers of monitor units for each beam. Thus all of the segments for each beam will have the same linearity error. Although these errors are small, they are systematic in nature and it is therefore possible to offset them. One solution could be to alter the number of monitor units used in a treatment plan according to the non-linearity of the dose output. Altering numbers of monitor units by 1% requires non-integer numbers. The latest Siemens Oncor linacs (along with some Elekta and Varian models) can deliver radiation with 0.1 MU precision. On beams of 6 or 7 MUs, this affords approximately 1.5% precision in altering the number, or around 1% for fields of 10 MU or more. While this precision is not good enough to completely eliminate non-linearity errors, it would allow any errors to be kept below 1%. A variable division scheme of the fluence profile as proposed by Beavis (2001) could also help, as it would mean that not all segments for each beam would use the same number of monitor units. However, as there are ranges of monitor units where the dose linearity is always above or below unity, the variable division scheme would have to try to distribute the segment sizes in such a way as to avoid too many similar-sized segments

that for which the errors would persist. It is difficult to quantify how much a plan would be affected by variable division, although it is likely that the non-linearity effect would be reduced, albeit in an unpredictable fashion. Figures 4.27 to 4.30 show that Patient H's plan has smaller dose errors from non-linearity than dose Patient S's plan. This is mainly because Patient S's plan uses segments of lower dose, and as shown in figures 4.7a/b, the dose non-linearity tends to be smaller for segments over about 5 MU. Thus not only do higher-dose segments have smaller errors, but modification of the monitor units used to compensate for those errors is easier. The easiest way to reduce non-linearity errors is therefore to increase segment size, by having less modulation and fewer segments. The errors, even with the small segments used in Patient S's plan, are small. However, they are systematic, and if added to overdosing errors from gravity motion and possible MLC miscalibration it will contribute to overall dosimetric errors in patient treatments.

The work here looks primarily at gravity-induced errors and therefore applies most readily to Siemens linacs and their double-focussed leaf design. The general principles of small MLC errors leading to large dosimetric errors are however applicable to all linacs. Elekta and Varian linacs utilise more reliable methods of positioning the MLC leaves during treatment, and the errors in the positioning are much smaller than those seen on the Siemens machines tested. As such, with errors generally below the 0.5 mm level, the graphs presented above show that the dosimetric errors likely for Varian and Elekta machines are small and most likely will not be noticeable for typical IMRT treatments.

## Conclusions

The work presented here shows that small systematic changes in linac behaviour can lead to large dosimetric errors in the delivery of step-and-shoot IMRT treatments. The sizes of the errors and the locations of hot and cold areas in the dose distributions are dependent upon the beam arrangement and specifics of the target site in the body. Incorrect or inaccurate calibration of MLC leaves could potentially lead to large dose delivery errors. Even with accurate methods of calibrating MLCs, engineering tolerances within the linac itself can lead to systematic changes in MLC position, characterised by gravity-induced movements within the backlash of the mechanical systems. These systematic changes lead to smaller dose errors than do poorly calibrated MLCs, although the dose errors could still be significant.

Choice of segmentation algorithm does not affect the dose errors caused by misbehaving MLC leaves, although there are small differences between some of the most common segmentation algorithms. The choice of algorithm can therefore be made based on treatment time or total MU considerations without significant effect on the dose distribution that results from the treatment.

The effects of gravity motion can be partly compensated for by deliberately miscalibrating the MLC leaves such that the gravity motion causes the MLC leaves to move from slightly too small to slightly too large. In this way, dosimetric errors in the dose distribution are shifted to be either side of 0, rather than always leading to hotspots. The dose differences can therefore be halved, leading to similar changes in analysis indices such as Kappa. The percentage of a target that fails the Kappa test is effectively threshold-based and halving dosimetric errors can all but remove failure of the Kappa index.

Systematic changes in linac output, such as a low-MU change in output, can also lead to significant dosimetric changes in treatments. There is no easy way to avoid this

during planning if inverse planning techniques are used. However, if the output curve is known then software can be used to alter the numbers of monitor units that treatment beams use to correct for the change caused by non-linear machine output. Clearly changes such as these require the use of non-integer monitor units for IMRT segments, which some centres may not have access to.

The work presented here has shown that an independent in-house dose calculation engine can prove invaluable in predicting the possible dosimetric errors caused by systematic changes that are too small to be incorporated into commercial treatment planning systems. Such software can easily be modified to include almost any kind of change in treatment machine behaviour, including both systematic and random errors. Future work could look at how random changes in linac behaviour, especially those affecting MLC placement, affect treatment plans. Such random errors would produce a blurred out dose distribution rather than the clear-cut cases of over or under dosing seen here, and an in-depth analysis of treatment margins and movements during treatment would have to be undertaken to analyse how such errors affect patient treatments.

This work deals solely with the physical aspects of dose distributions, namely the dose and location of that dose. Suggestions and methods have been laid out for reducing the dose errors and improving the accuracy of IMRT as delivered using the equipment available during the course of this work. Improving patient outcomes is the ultimate goal of radiotherapy research, and there is clear scope for analysing how the problems looked at here will affect patient outcomes and cure rates. There is a great deal of work to be done in radiobiology, and chapter 6 attempts to touch upon some of it and deal with how the dosimetric errors simulated and investigated here will impinge upon patient care.

# Chapter 5 – Is there any equivalence of dose difference and distance-to-agreement in dose evaluation metrics?

## Introduction

Differences in dose between what is planned for a treatment and what is delivered to a patient are relatively easy to measure in their physical form with the use of radiographic film and ionisation chambers to check the dose on or near a patient or inside a test phantom. Relating dose difference to an effect is more difficult as an overdose in one area may not balance out an underdose in another area. The mean dose over a volume could be that which was planned, but still have significant deviation which is likely to have some sort of effect on the outcome of a treatment. Thus the analysis of dose difference or distance to agreement is limited in its usefulness and the dose differences need to be extrapolated to another level in order for them to become useful in assessing what is and isn't a good dose distribution.

The equivalent uniform dose or EUD (Niemierko 1997) attempts to reduce a heterogeneous dose distribution across a tumour to a single number, the dose which if applied uniformly across the tumour would produce the same biological end-point as the heterogeneous dose. The original formulation of EUD is shown in equation 1 below:

$$EUD = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp\left(-\frac{(S-\tilde{S})^2}{2\sigma^2}\right)\bullet EUD(S)dS$$

where:

$$S = \ln[-\ln(SF_2)],\ \tilde{S} = \ln[-\ln(\tilde{S}\tilde{F}_2)] \qquad (1)$$

$$EUD(S) = \frac{N_f}{D_{ref}}\left[-\frac{\alpha}{\beta}+\sqrt{\left(\frac{\alpha}{\beta}\right)^2+4\frac{D_{ref}}{N_f}\left(\frac{\alpha}{\beta}+D_{ref}\right)\frac{\ln A}{\ln(SF_2)}}\right]$$

$$A = \sum_{i=1}^{N} V_i\rho_i(SF_2)^{\frac{D_i}{D_{ref}}\frac{\alpha/\beta+D_i/N_f}{\alpha/\beta+D_{ref}}} \Bigg/ \sum_{i=1}^{N} V_i\rho_i$$

The equation depends on many factors, including: $SF_2$, the surviving fraction of cells at the $D_{ref}$ of 2Gy; $\overline{SF_2}$, the average $SF_2$ for a population; the clonogen density $\rho$; and $\alpha$ and $\beta$, the dose response values for tumours. The formula attempts to include almost all factors that can affect the response of a tumour to radiation, and is based on probabilities of certain proportions of cancerous cells surviving irradiation. Factors such as $\alpha$ and $\beta$ relate to rates of single and double strand DNA breaks in a cell during a single irradiation, and thus relate to the effect of fractionation as a proportion of single-strand breaks will be repaired by the cell between fractions. The equation takes into account partial irradiation of volumes and population spread of radiosensitivity. Many of these factors cannot be derived accurately, and the equation, although based on good scientific theory, is difficult to use for quantitative radiobiology.

In 1999, Niemierko introduced the generalised EUD equation as an approximation to the more complicated formula presented two years earlier:

$$EUD = \left( \frac{1}{N} \sum D_i^a \right)^{1/a} \qquad (2)$$

where $D_i$ is the dose to a voxel $i$ of the structure, $N$ is the total number of voxels and $a$ is the tissue-dependent factor. The equation is based on the power-law response of many biological systems, and can be derived from the more complicated EUD equation with suitable assumptions and simplifications (Li *et al* 2003). The generalized EUD can also be applied to organs at risk, being mathematically equivalent to the methods used by Burman *et al* (1991) in the calculation of normal tissue complication probability (NTCP). The generalized EUD equation itself is known as a generalised average (Abramowitz and Stegun 1968) and is thus a good mathematical way of showing the equivalent uniform dose. This work will use the generalized EUD equation, which will be referred to simply as EUD from this point forward.

The clear advantage of this formula is its simplicity, the ease of implementing the function as computer code and the inclusion of only one tissue-dependent factor, $a$, which controls how the formula treats inhomogeneous dose within the distribution that is analysed. Although the formula is much simpler than the full mechanistic EUD equation, Wu *et al* (2002) show that it is a good fit to phenomenological data. With $a$=1, the formula calculates the arithmetic mean of the dose distribution, whereas $a$=∞ and $a$=-∞ return the maximum and minimum doses, respectively. It is clear therefore, that: highly serial organs at risk, for which the maximum dose received represents the overall damage caused, require high positive values of $a$; more parallel OARs require smaller positive values of $a$; and tumours, where the minimum dose becomes an important factor in the survival of the tumour, require negative values of $a$. Very aggressive tumours require more negative values of $a$ than do more radiosensitive ones.

The EUD calculation is a useful tool in assessing radiobiological effects of non-ideal dose distributions to both normal tissues and target volumes in radiotherapy – an oncologist can decide if the inevitably inhomogeneous dose distribution across a target presents an acceptable EUD from which the likely tumour control probability can be calculated.

Patient-specific quality control of IMRT plans typically involves measuring a dose distribution and comparing it to that calculated by the treatment planning system (TPS). Such comparisons are increasingly relying on quantitative measures of the congruence of the reference dose distribution (calculated by the treatment planning system) against the evaluated dose distribution (measured or simulated as a quality check, see Low 2001 for definitions). Several of these quantitative measures are discussed in-depth in chapter 3. In such evaluations, for example the Gamma index (Low *et al* 1998), the concept of dose difference and distance-to-agreement tolerances (see chapter 3) are required as inputs to the evaluations. Typically these tolerances are

set based on feelings of what the prescribing oncologist will allow as errors in the dose distribution, or what is decided is achievable by the verification physicist. There is a tendency in the literature (Stasi *et al* 2005, ICRU 1987, Agazaryan 2003) to set the same DTA tolerance in millimetres as the percentage dose difference tolerance, for example 3%/3 mm, 4%/4 mm and so on. The reasons for setting the same values for each have never been made clear, and the equivalence of % dose difference to mm DTA has never been shown.

The EUD provides us with a quantitative radiobiological view on dose distributions and errors within them. By comparing the different EUDs of a reference and evaluated dose distribution, we can make more informed decisions about dose and distance tolerances for future QA work. This study aims to use the EUD to determine the equivalence or otherwise of % dose difference and millimetres DTA for dose distributions in terms of their usage within dose evaluation metrics.

## Materials and methods

### 1) Dose distributions and phantom

All dose distributions used were taken from an in-house treatment planning system, ARPS version 1.90. The planning system has been separately commissioned and has been in clinical use for conformal radiotherapy for over 5 years. Two different dose distributions were used for the tests. The first used 13 equally-spaced coplanar beams to create a dose distribution that was as close to uniform as possible across the target volume and circularly symmetric to match the target volumes chosen. This beam arrangement provided relatively shallow dose gradients at the edge of the target volume in the transverse direction, and steep gradients in the perpendicular direction, along the axes of the cylinders.

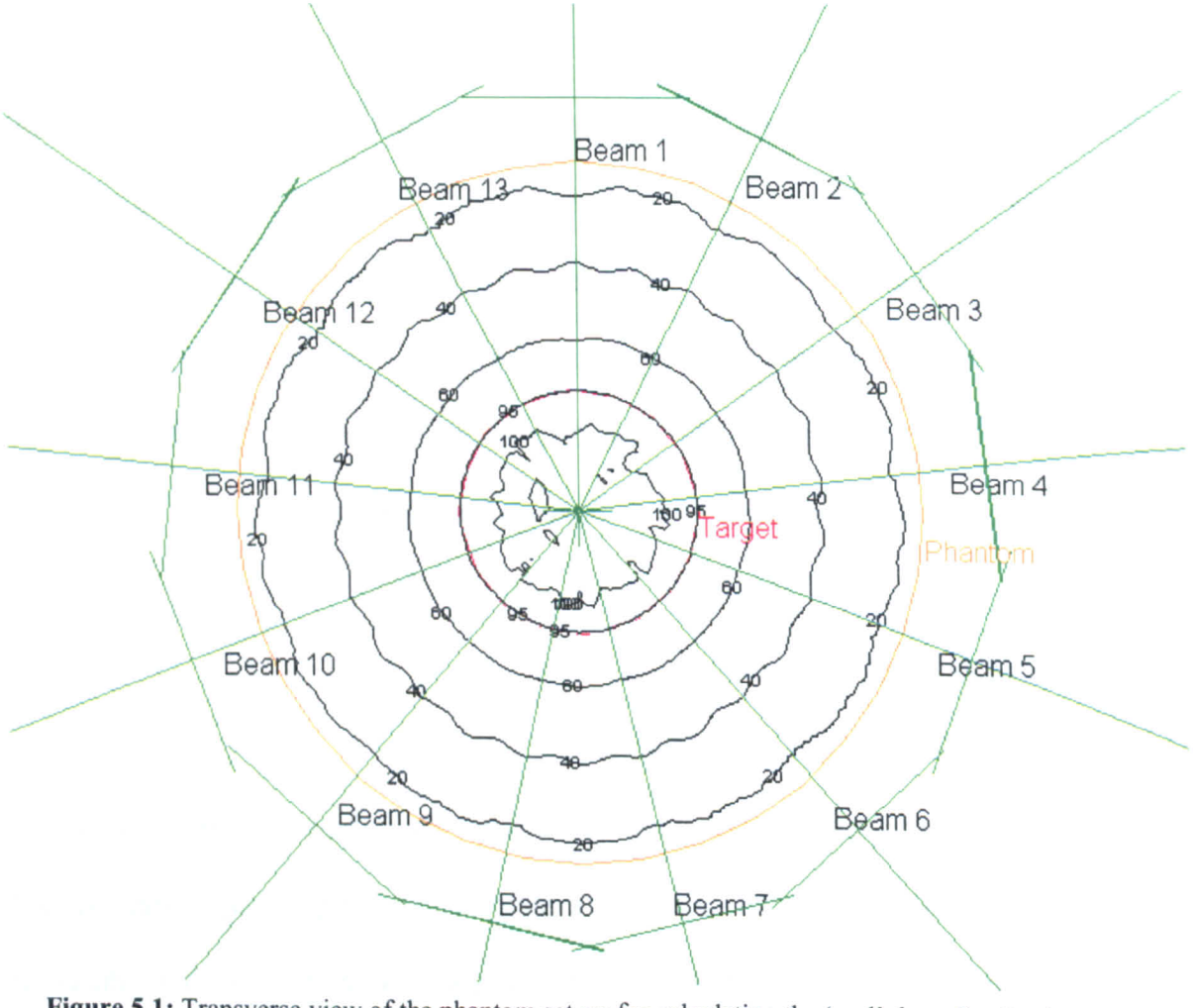



**Figure 5.1:** Transverse view of the phantom set-up for calculating the 'real' dose distributions. The phantom is cylindrical, with a diameter of 200 mm, and the target shown is a 70 mm cylinder. The 13 incident beams are shown, where the perpendicular lines indicate the field widths. At the centre of the phantom, a circularly symmetric dose distribution is created and the 95% isodose (black) follows the edge of the target (red) closely.

The second dose distribution used was taken from a 3-field clinical plan for a prostate cancer patient. The plan was selected from the database of current patients on treatment at our centre as a typical example of a 3-field pelvic plan. The three beams are oriented from the anterior, left and right sides of the patient and are shaped using Siemens MLCs. The starting point for the plan was to shape the MLCs to be 8 mm larger than the PTV border. The lateral beams were wedged and weighted to achieve a near-uniform dose in the target volume, and then individual MLC leaves were moved to improve coverage of the 95% isodose around the PTV. The dose was normalised to 100% at the isocentre, and the 95% isodose in the isocentric transverse plane is approximately square with rounded corners.

**2) Target volume and patient structures**

For the circularly symmetric dose distributions, the 'patient' was a 20 cm diameter cylinder of unit density created within the treatment planning system. The cylinder was 30 cm in length with its axis aligned perpendicular to the transverse plane such that each transverse slice was circular. This shape best allows the incident beams to create a circularly symmetric dose distribution at the centre of the patient. The target volumes were also cylindrical, placed at the centre of the patient cylinder aligned with its axis. The target volumes were used as a planning tool only – the incident beams used were rectangular and sized in order to conform the 95% isodose to the cylinder's edges as well as possible. The resulting 95% isodose was then taken as the target volume's new edge. The resultant shapes were cylinders with slightly rounded ends, which approximate real patient targets better than cylinders with right-angled ends. The situation also models the ideal scenario in which the 95% isodose conforms perfectly to a target's edges. Targets of various lengths and diameters were used, to investigate how the changes in EUD behave as a function of target size. Figure 5.1 shows the beam

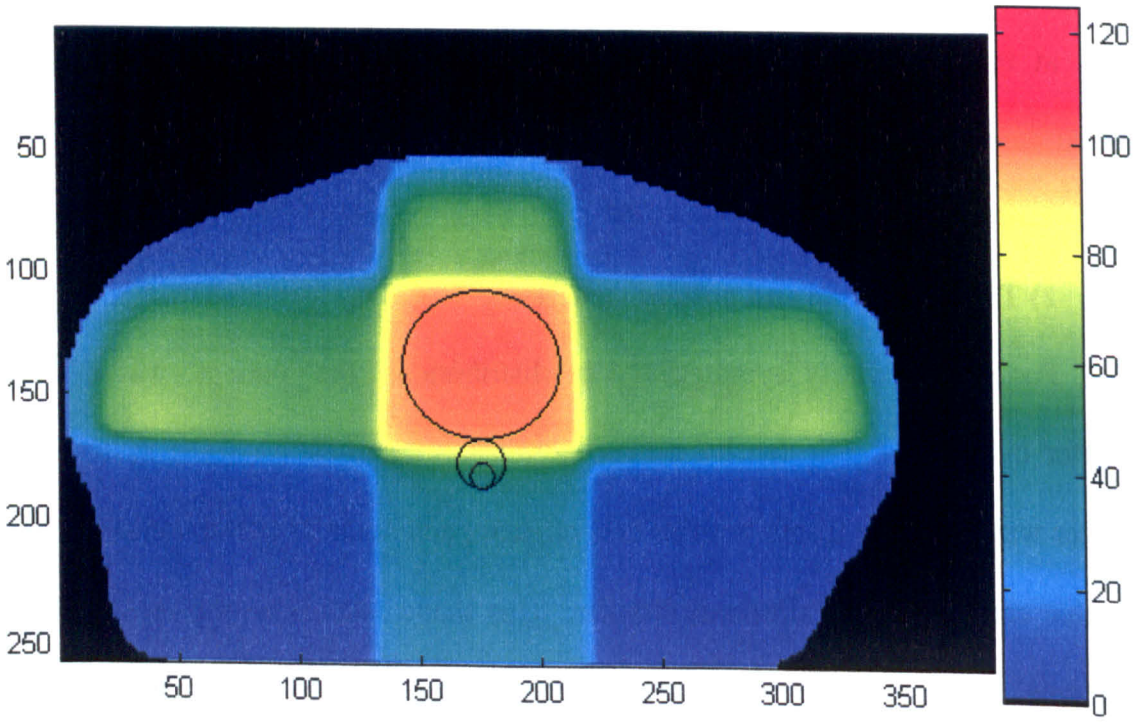



**Figure 5.2:** Dose distribution used from the 3-field plan, showing the 3 beams from the anterior (top), patient left (right) and patient left (right). The view is a transverse slice through the patient looking towards the patient's head. The colour scale is defined in percent of the isocentric dose. Shown as black contours are the outlines of the target (large ellipse) and the 'rectum' and 'spine' structures (medium and small circles respectively).

arrangement for this set-up, including a 70 mm diameter target and the calculated isodose lines.

For the real 3-field plan, the dose plane from the isocentric transverse slice was selected and the target volume was approximated as the largest ellipse that would fit inside the 95% isodose, with the 100% dose point defined at the isocentre. This ellipse had its semi-major and semi-minor axes aligned in the left-right and posterior-anterior directions respectively, the horizontal radius being 33 mm and the vertical radius 32 mm. To investigate the effects of spatial movements of structures that are not target volumes, a 'rectum' was added, modelled as a cylinder of diameter 20 mm, located posterior to the target with its anterior edge just touching the posterior edge of the target. In the single transverse slice this 'rectum' became a 20 mm circle. For a second test of normal structure effects, the target was unchanged and a 'spine' was created as a 10 mm circle posterior to the target with its anterior edge 10 mm from the posterior edge of the target. Figure 5.2 shows the dose distribution for the slice used and the outline of the structures added.

**3) EUD calculations**

The EUD formula was implemented in MATLAB along with the code to move target volumes relative to the dose distributions. The dose voxels within the targets were extracted for each dose movement and the EUD was calculated.

An initial test was carried out using a target volume 80 mm in diameter and 80 mm long. This volume was moved within the circular 13-beam dose distribution in the transverse plane. The EUD of the dose distribution within the cylinder was calculated for various values of the EUD parameter $a$. Values used ranged from minus infinity to +5. The value of $a$ to use for all subsequent calculations was chosen based on this first

| | Target shape | Target size | Dose distribution | $a$ | Directions moved | Figures |
|---|---|---|---|---|---|---|
| 1 | Cylinder | Length 80 mm<br>Ø 80 mm | 13-beam circular | Various<br>$-\infty$ to +5 | In-plane | 5.3, 5.4 |
| 2 | Cylinder | Length $\infty$<br>Ø 17-110 mm | 13-beam circular | -15 | In-plane | 5.5 |
| 3 | Cylinder | Length 13-82 mm<br>Ø 26 mm | 13-beam circular | -15 | Along axis | 5.6 |
| 4 | Cylinder | Length 13-82 mm<br>Ø 70 mm | 13-beam circular | -15 | Along axis | 5.7 |
| 5 | Cylinder + 20 mm 'rectum' | Length $\infty$<br>Ø 65 mm | Real, 3-field plan | Target: -15<br>'Rectum': +8 | In-plane | 5.8, 5.9 |
| 6 | Cylinder + 10 mm 'spine' | Length $\infty$<br>Ø 65 mm | Real, 3-field plan | Target: -15<br>'Spine': +20 | In-plane | 5.8, 5.9 |
| 7 | Cylinder + 'rectum' and 'spine' | Length $\infty$<br>Ø 65 mm | Real, 3-field plan | Target: -15<br>'Spine': +20<br>'Rectum': +8 | Obliquely in-plane | 5.10 |

**Table 5.I:** Summary of target shapes and sizes, dose distributions and parameters used for the EUD tests. In-plane movements are in the transverse plane.

set of calculations. Andrzej Niemierko suggests[1] using $a$=-15 as a good mid-range value for relative calculations, which agreed with the data (figure 5.3). It is possible to fit the EUD equation 2 to the clinical data reported by Terahara *et al* (1999) to attain a value for $a$ for chordoma local control of -13, while Chapet *et al* (2005) suggest that $a$=-5 would correspond to a radiosensitive tumour and $a$=-20 to an aggressive tumour. Based on this, a value of $a$=-15 was chosen for subsequent calculations.

For these subsequent tests, cylinders of various lengths and diameters were moved laterally and longitudinally within the dose distributions using $a$=-15 to analyze the effects of size on the changes in EUD. Table 5.I summarizes all the experiments performed. Cylinders moved laterally in-plane (perpendicular to their axis) were assumed to be infinitely long, as the effects of a slightly different dose at the ends of the cylinders were found to be negligible. Thus the EUD was calculated by analysing a single slice of the dose data and performing the movements within the plane. Longitudinal movements used the full 3D dose distribution calculated by the treatment planning system.

For the introduction of 'rectum' and 'spine' structures, different values of $a$ were required. Li *et al* (2003) show that the value of $a$ is equal to $1/n$ for the values of $n$ reported by Burman *et al* (1991). Thus for the 'rectum' a value of $a$=8 was used, with $a$=20 for the 'spine' (both values rounded to the nearest integer).

**Results**

Figures 5.3 to 5.10 show the results of the changes in EUD as a function of spatial shift for several factors. Table 5.I lists which figures show which results.

[1] Personal communication

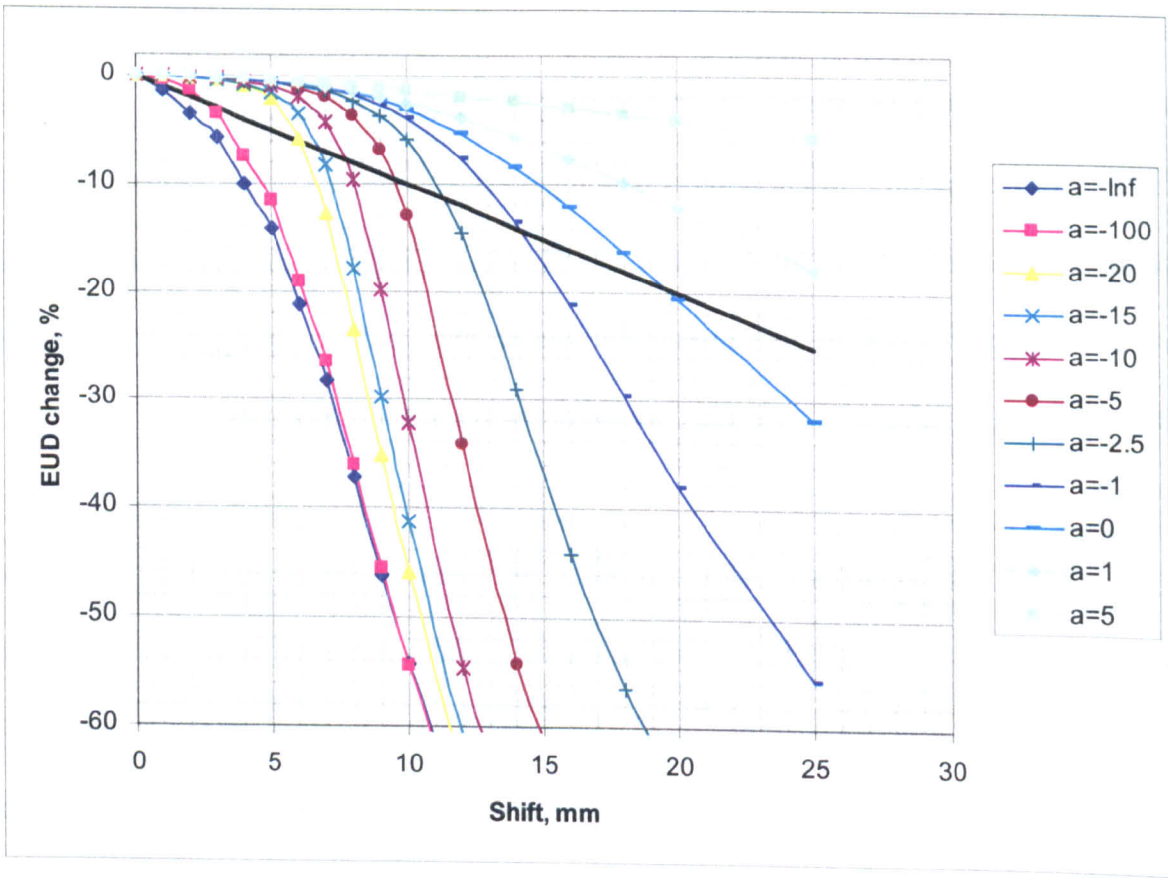



**Figure 5.3:** Change in EUD as a percentage of prescription dose as a function of EUD parameter *a* and shift of the target structure. The target is cylindrical, in an idealised cylindrical dose distribution. All movements were performed in the direction of the cylinder's axis.

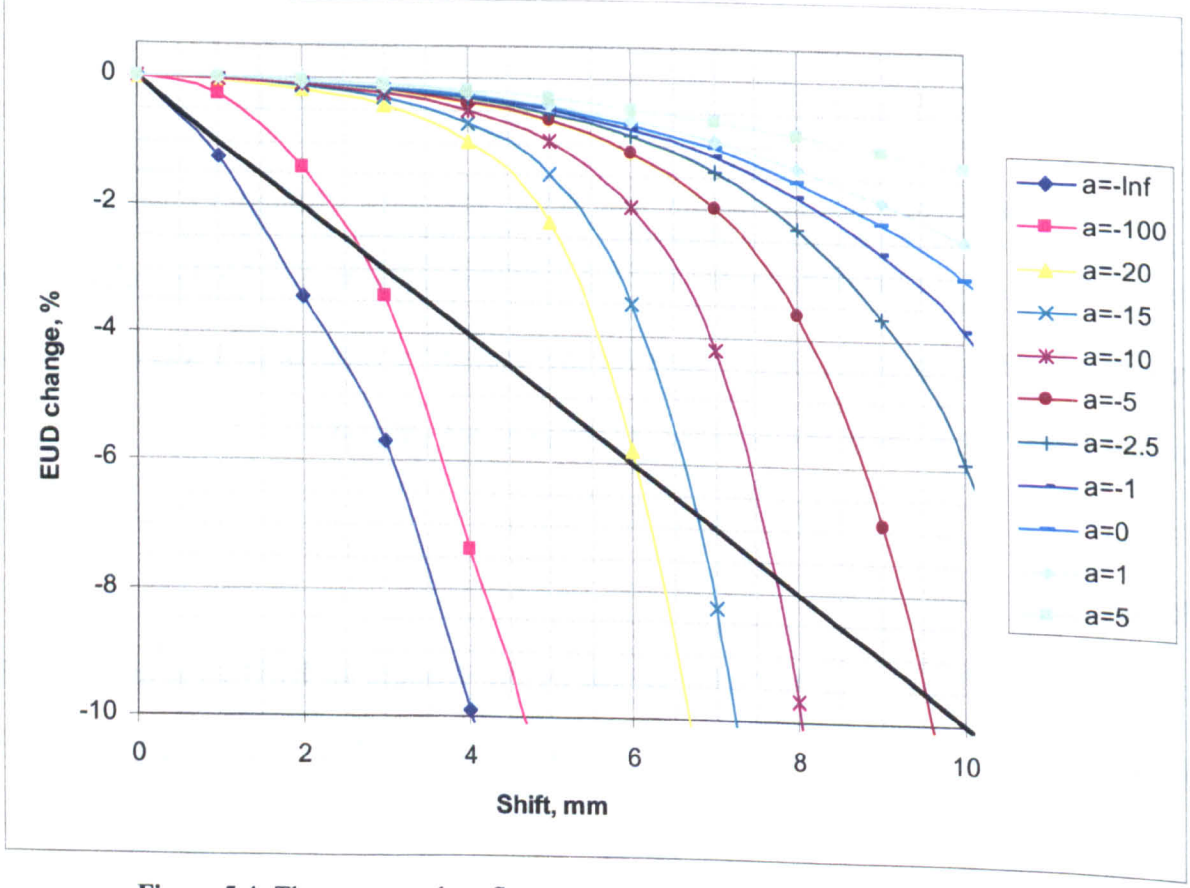



**Figure 5.4:** The same graph as figure 5.3, zoomed to show the upper-left portion.

The graphs show the change in EUD as a percentage of the prescription dose to the target as a function of the movement of the target. Thus if the prescription dose is 60 Gy, a 4% indicated change in EUD is a 2.4 Gy change in equivalent dose. The actual EUD in Gray for 0 mm movement of the structures varies slightly according to the size of the target and the value of *a* because although the dose distribution was made as uniform as possible, there remain small variations in the dose across the target volume. These small changes varied depending on target size, as the beam penumbra widths are a constant and therefore affect smaller and larger volumes differently, and also on the value of the power factor *a* in the EUD formula.

For all graphs, the solid black lines show the locus of 1% change in EUD per millimetre movement. This line represents the equivalence of percentage dose change and distance-to-agreement numerical values such as 3%/3 mm.

**Variation of parameter 'a'**

Figure 5.3 shows the general shape of the EUD curves for a spatial shift of the target structure of up to 25 mm. Movements of this magnitude are not clinically relevant, so figure 5.4 shows a zoomed-in view of shifts up to 10 mm. The target used for these calculations was an 80 mm diameter cylinder, moving in the transverse plane, with the different lines showing different values of *a* used in the EUD calculation. The curves show that for values of *a* greater than -20, the EUD changes slowly at first and then more rapidly as the spatial shift becomes larger. For $a > -20$, all EUD change curves lie above the 1%/mm equivalence line for shifts smaller than 6 mm. Taking the $a = -15$ curve as an example, the graph shows that a 6 mm movement of the target structure causes a change in EUD of only 3.4%, a movement of 4 mm causes a change of only 1% EUD. As the value of *a* increases towards 0, the curves become shallower such that at the $a = -5$ level, a 6 mm movement causes a change in EUD of just 1.2%.

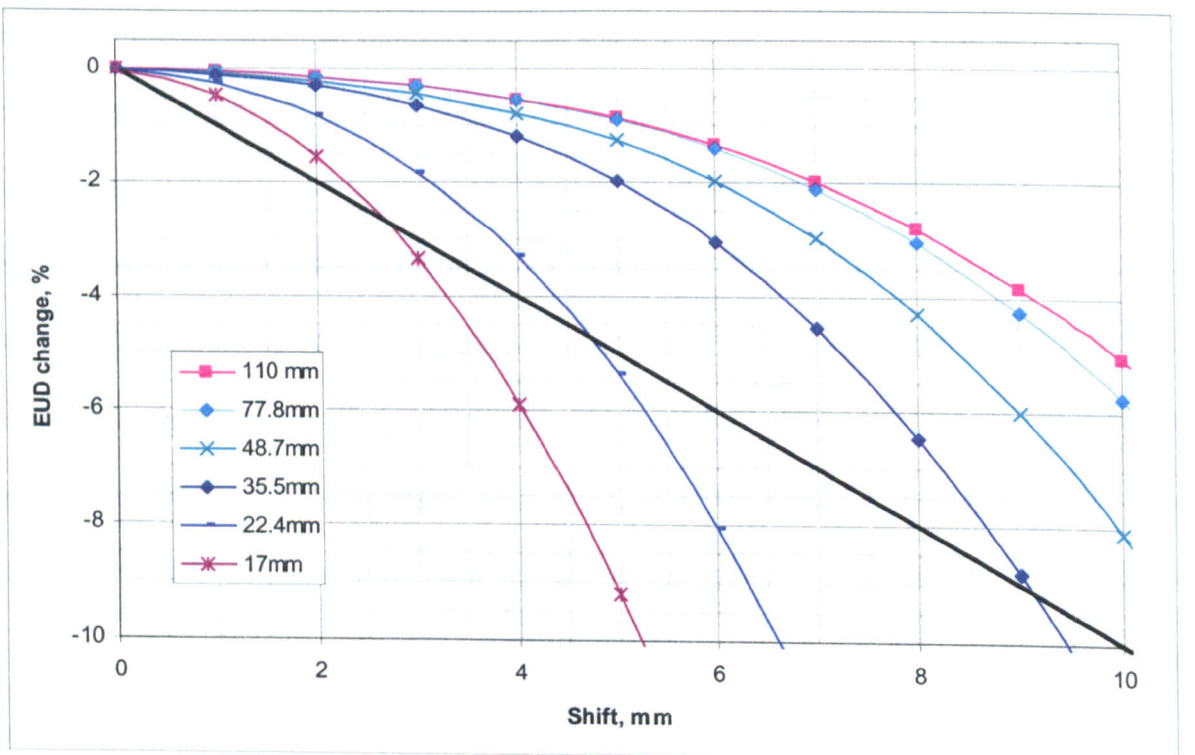



**Figure 5.5:** Change in EUD as a percentage of prescription dose for cylinders with $a$=-15. Cylinders are moved perpendicular to their axis, and are of varying diameters as shown. The dose distribution is a realistic circularly symmetric delivery from 13 beams.

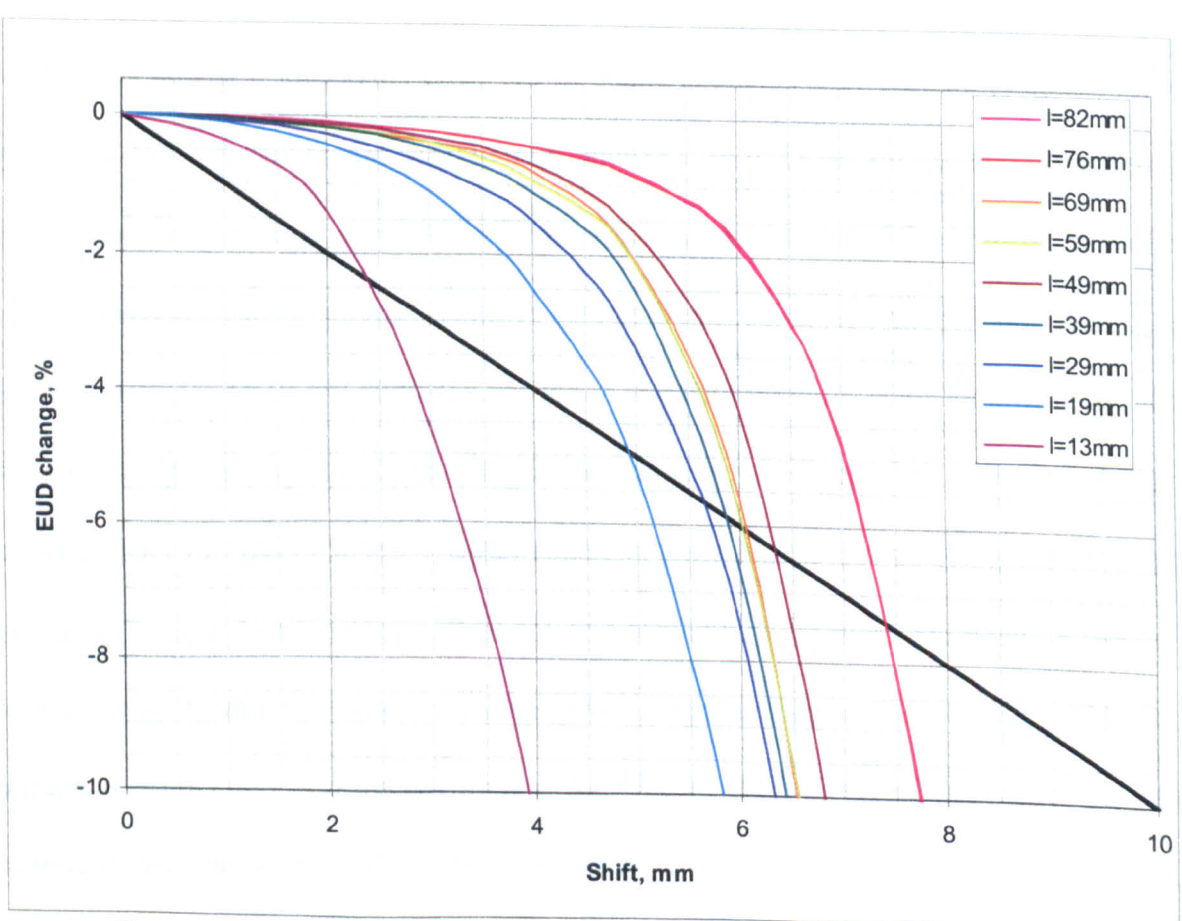



**Figure 5.6:** Change in EUD as a percentage of prescription dose for 26 mm diameter cylindrical targets moving along their axis. The cylinders are of varying lengths, as shown, and the EUD was calculated using $a$=-15

**Variation of target length**

Figure 5.5 shows the change in EUD as a function of varying target diameter, for $a$ = -15 and in-plane movement of the target. The diameters chosen represent the typical range of tumour volumes treated – cylinders smaller than 17 mm were not investigated as the calculation accuracy of the treatment planning system becomes worse with the very narrow fields involved. As expected, the curves show that small volumes have a steeper change in EUD than do larger volumes. The same movement for a small structure moves more of it into the penumbral regions around the target area more quickly than for a larger volume. The curves show that for all but the smallest cylinders, the change in EUD for movements up to 5 mm is smaller than 5%, with all the curves lying above the black equivalence line in this region of the graph. For the largest structure, 110 mm in diameter, a 7 mm movement causes a change in EUD of just 2% - parity between distance moved in millimetres and % change is not reached until a movement of 18.5 mm.

**Variation of target length**

Figure 5.6 shows the changes in EUD for a 26 mm diameter cylinder of various lengths, as it is moved along it axis. The parameter $a$ is -15 for all of the curves. As with the previous graphs, the curves all lie above the black equivalence line for small shifts, with the smallest volumes falling below the line first. The curves again show a relatively flat profile for small movements of the structures, followed by a rapid steepening. The curves fall more quickly for large movements here because of the direction of movement. The incident radiation beams are coplanar and are perpendicular to the direction of motion for this test. Thus in the direction of movement, the dose penumbra is created by the beam edges and is therefore as steep as a single beam's edge penumbra. The dose falls very quickly to about 10% as the only

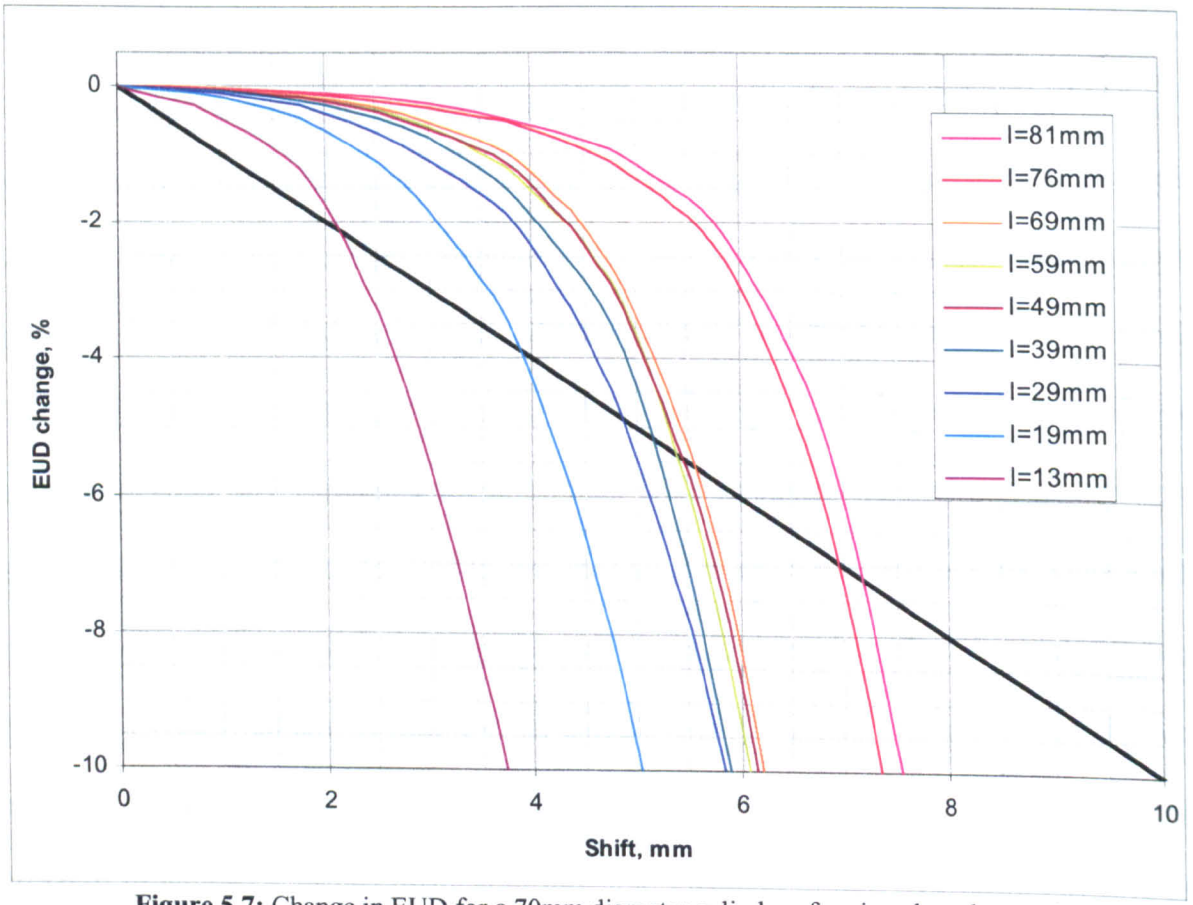



**Figure 5.7:** Change in EUD for a 70mm diameter cylinder of various lengths moving along its axis. The EUD parameter was $a$=-15.

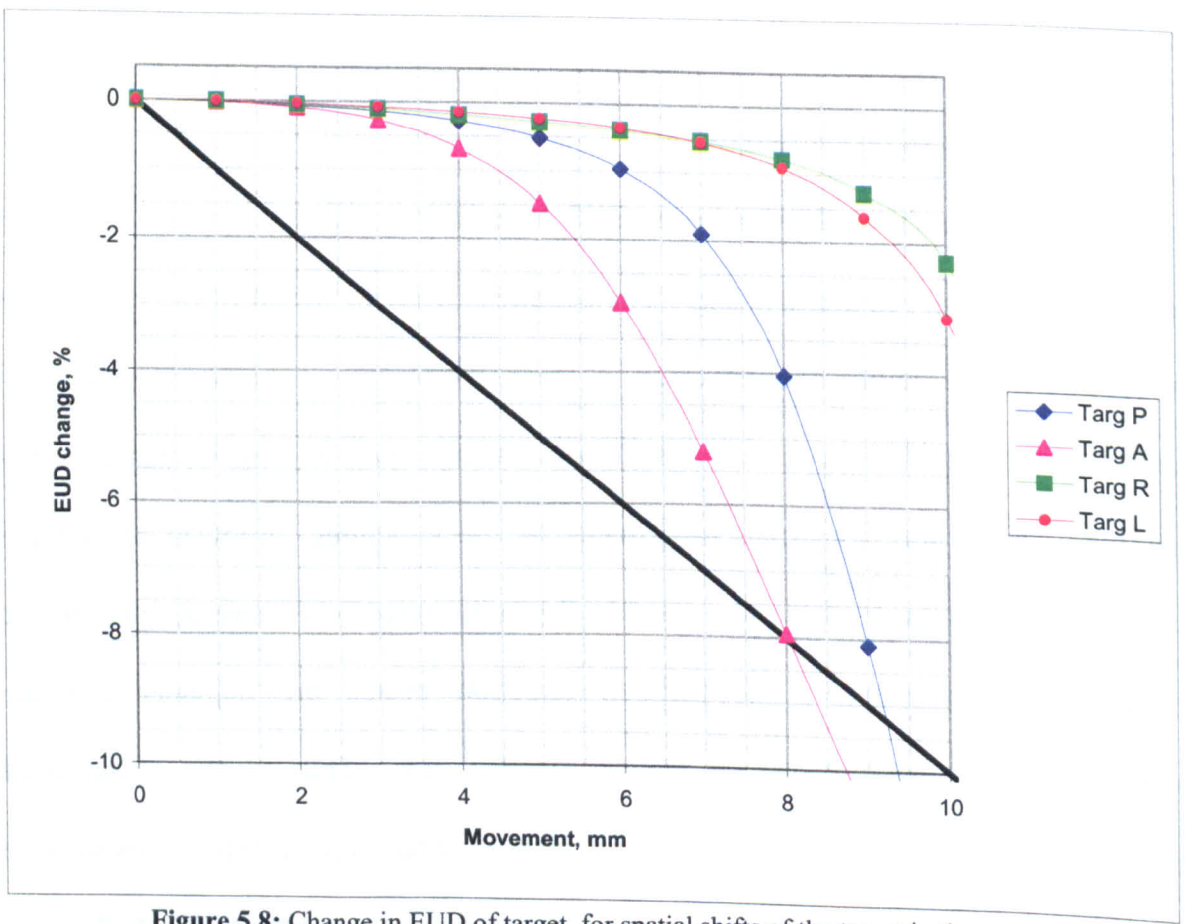



**Figure 5.8:** Change in EUD of target, for spatial shifts of the target in the Posterior, Anterior, Left and Right directions. Dose distribution is taken from a real 3-field treatment plan, and an EUD parameter of $a$=-15 is used.

dose delivered in the regions superior and inferior to the target is from scatter. Therefore as the target structure is moved along its axis, the dose distribution inside it becomes highly inhomogeneous more quickly than for the in-plane movement shown in figure 5.5 and the EUD falls more quickly as a result. However, for structures between 19 and 82 mm in length, the parity between EUD change and distance occurs between 5 and 7 mm, which is typically larger than the DTA tolerances used for dose verification. Structures longer than 82 mm will clearly show dose/distance parity at larger values.

Figure 5.7 shows the changes in EUD for a 70 mm diameter cylinder of various lengths, moving along its axis. The set-up here is the same as for the results shown in figure 5.6, with a larger diameter cylinder. The curves are very similar to those for the 26 mm cylinder. This is to be expected, as the fractional volume of a cylinder along its length is independent of diameter. The small changes in the EUD curves are due to the slightly different rounded dose distributions at the end of the differently sized cylinders. For cylinders of lengths 19 to 81 mm, the parity between dose change and distance moved occurs between 4 and 7 mm.

**Movement of simulated clinical structures**

Figures 5.8 and 5.9 (over page) show the results of moving the 'tumour', 'rectum' and 'spine' structures in-plane in a 3-field clinical dose distribution. The curves for the target in figure 5.8 follow a similar shape to those presented for the cylinders in the previous figures, as expected. The small differences in the change in EUD between the target moving left and right are caused by the slight asymmetries in the original dose distribution, as the patient and beam arrangement are not themselves symmetric. The curves for left and right movements of the 'spine' and 'rectum' in figure 5.9 are almost flat, showing no change in EUD for movements up to 15 mm. This is because the structures are small compared to the diameter of the 'tumour' and

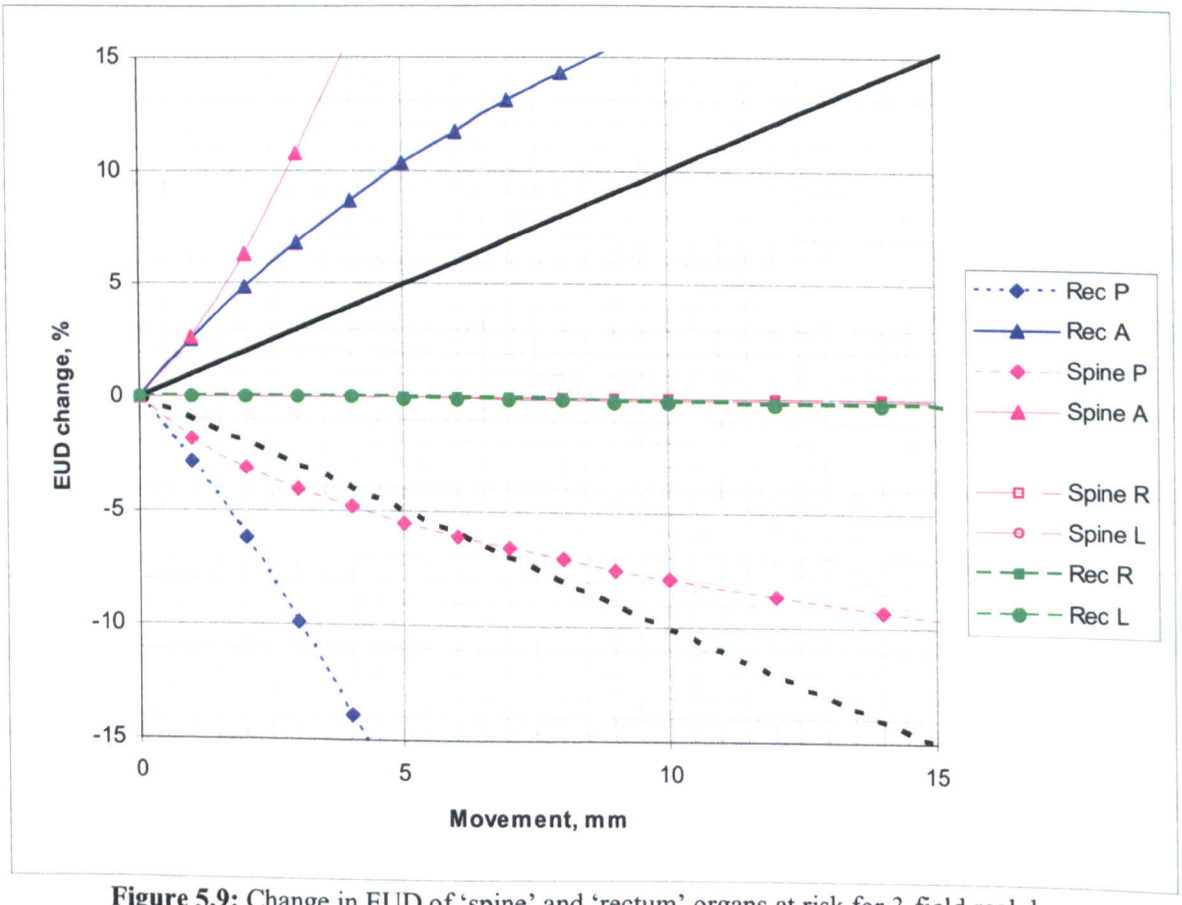



**Figure 5.9:** Change in EUD of 'spine' and 'rectum' organs at risk for 3-field real dose distribution. EUD parameters used are spine=+20, rectum=+8. EUD changes are shown for movements of the structures in the posterior, anterior, left and right directions. Right and left movements of the spine and rectum overlay each other along the 0% line.

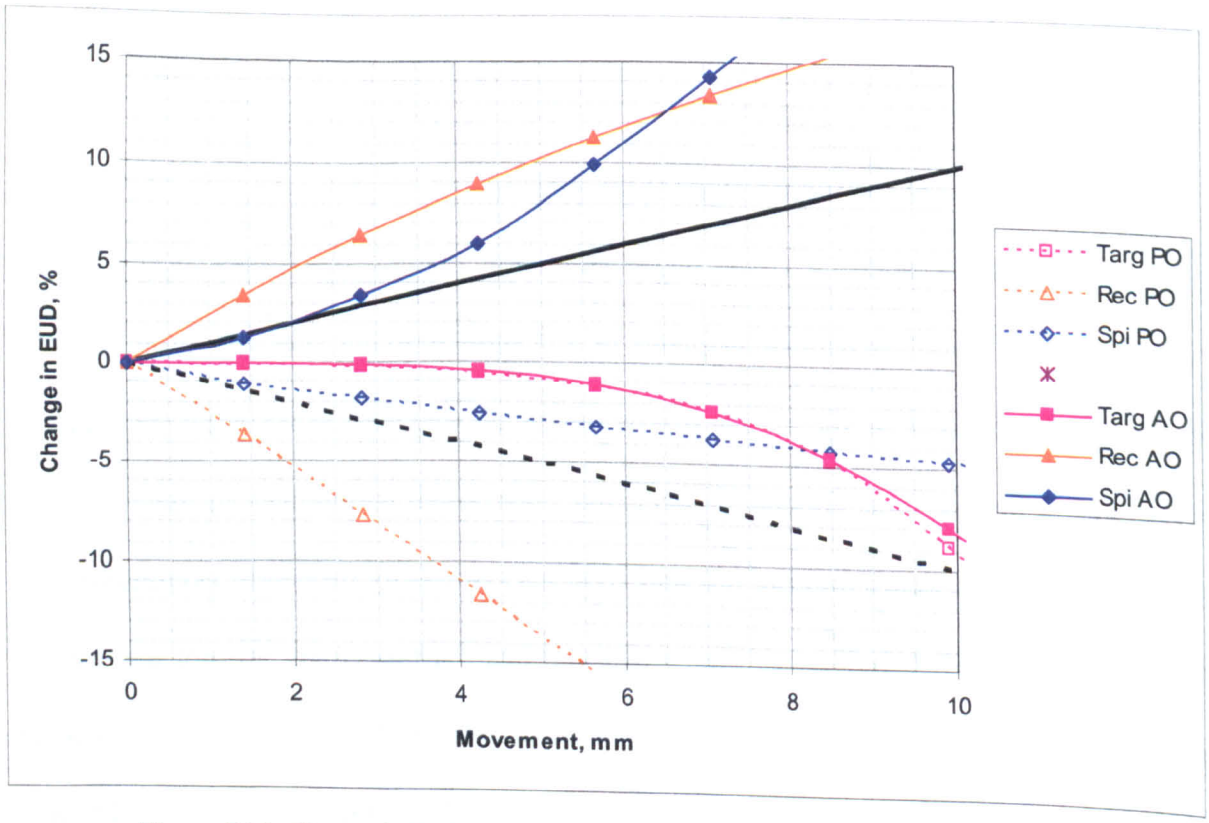



**Figure 5.10:** Change in EUD of 'spine', 'rectum' and target structures for real dose distribution and oblique movements of the structures relative to the dose. PO represents a 45° posterior oblique movement, AO represent a 45° anterior oblique movement. The values of $a$ are as those for figure 5.9

are located in the path of the anterior beam. As the structures move left and right they remain within the beam and the dose variation along that direction is low until the edges of the beam are encountered after a movement of approximately 25 mm.

As with all previous figures, the curves for the movements of the target (figure 5.8) show that the change in EUD in % is smaller than the movement of the target in mm for movements up to 8 mm. This value of 8 mm is slightly higher than for the cylinders presented above due to the beam arrangement used. The 13-beam arrangement used for the cylindrical volumes provides a uniform, relatively steep dose gradient in all directions. The use of 3 perpendicular beams for the clinical case provides much shallower and more realistic dose gradients in the 4 cardinal directions as a movement in any of these directions keeps the target within the paths of at least 1/3 of the incident beams. Thus for clinical cases, the parity between dose difference and distance moved occurs at higher values than for the more ideal dose distribution presented previously.

Figure 5.9 shows curves for the 'rectum' and 'spine' that are outside of the 1%/mm equivalence line for even the smallest of structure movements. For posterior movements of the 'spine' and 'rectum' structures relative to the dose distribution, the change in EUD is negative. This is expected, as the structures are moving away from the anterior beam and further from the edges of the lateral beams. For both structures, the percentage change in EUD is greater than the movement in mm, both curves lying below the black equivalence line for movements of less than 6 mm. For movements larger than this, the 'spine' structure crosses the equivalence line, as the structure is in the low-dose region of the exit dose of the anterior beam. The EUD for this structure starts at 50% of the target dose. The dose distribution in the posterior direction of the target falls slowly to approximately 30%, hence the slow fall of the 'spine' EUD. The EUD for the 'rectum' structure falls much faster because the EUD starts at

approximately 77% and falls to a similar level as the 'spine'. The 'rectum' starts off closer to the target and thus starts with the higher EUD, but moves quickly into the same area of dose as the 'spine'. Although negative changes in EUD for organs at risk are an unexpected bonus of changes in dose distributions, this analysis shows that there is no implicit equivalence between distance and percentage change in dose, even for negative changes in dose. Such negative changes merit analysis even though they are good for the patient as a negative change in the EUD of both target and organ at risk could actually result in an overall lower probability of uncomplicated control.

For anterior shifts of the 'spine' and 'rectum' structures, the EUD curves are again outside of the equivalence line, with a 5 mm movement equating to a 10% change in EUD for the 'rectum' and a 21% change in EUD for the 'spine'. The EUD of the 'rectum' structure starts at 77% and rises as the structure moves further into the high-dose area, although because it already has one edge in the high-dose area, with its anterior edge touching the 95% isodose, the rise is not as dramatic as for the 'spine'. The 'spine' structure begins at 50% EUD, but located just outside the edges of the lateral beams. Thus a small anterior movement of the structure moves it across the high gradient penumbra present there and rapidly increases the dose across the structure and correspondingly the dose.

Figure 5.10 (previous page) shows the effect of moving the structures diagonally within the dose distribution rather than in the four cardinal directions. As the dose distribution is approximately symmetric from left to right, only diagonally-anterior and diagonally-posterior directions are shown, as Post-Oblique and Anterior-Oblique. The target curve behaves as with many of the curves before, with little reduction in EUD for small movements followed by a steepening of the curve. The equivalence between distance and dose occurs at around 10-11 mm, which is between that for anterior-posterior and left-right movements, as expected. The organ-at-risk structure curves also

show properties between those of ant-post and left-right movements. The gradients of the curves are smaller, which bring them closer to the black equivalence lines shown. The curve showing the change in EUD for the 'spine' structure shows the best equivalence between distance and dose, with the post-oblique curve just inside equivalence and the ant-oblique just outside for spatial shifts less than 5 mm. The curves for the 'rectum' structure are both outside of the equivalence line for this structure, although the difference is not as great as for the ant-post movements shown in figure 5.9.

**Discussion**

All the calculations of EUD carried out show that, for realistic shifts and target sizes, the percentage change in EUD of the PTV is smaller than the movement of the structure in millimetres. For all target sizes and movement directions, this was only true for small movements, with the percentage change in EUD for very large movements always greater than the distance moved. The crossing point of the EUD curve and the 1%/mm equivalence line varied from 2 mm for the smallest target sizes and most extreme dose gradients (13 mm long cylinder, moving along axis) to 18.5 mm for larger volumes in shallower gradients (110 mm cylinder moving transversely). For all target sizes except the smallest studied, the crossing point was at 5 mm or more. This shows that for dose evaluation metrics that combine dose difference and distance-to-agreement, the use of the same number for the % difference and millimetre shift is not valid for analysis of target volumes. For a target volume of 70 mm diameter and 49 mm length, a movement of 4 mm causes a change in EUD of only 1.5%, and the change is smaller for larger volumes, typical of targets such as prostates – the change in EUD for a 70 mm cylinder of length 76 mm is only 0.7% for a 4 mm movement. These are results for movements in the direction of steepest dose gradient – for movements in

transverse planes, where dose gradients are much shallower for coplanar treatments, the EUD curves show a shallower falloff.

The EUD is a good indicator of the outcome for the patient of the radiotherapy treatment (Ebert 2000). The radiobiological aspects of it are important to consider when analysing differing dose distributions – it may be that the dose distributions agree within a certain dose and distance tolerance, but physical dose metrics such as the Gamma index are purely physical. They cannot give an indication of how the outcome of a treatment will change given a non-ideal dose distribution. EUD provides a simple tool for estimating the effects of heterogeneous dose distributions, especially those caused by variances in delivery of the dose distribution and differences from the treatment planning system.

If we apply the EUD curves to dosimetric verification of treatments and consider only target volumes, important implications for treatment planning arise. Treatment margins are included into planning to ensure that the CTV gets the target dose even when movements such as patient setup, internal organ motion and breathing are taken into account. However, if it is considered acceptable to have up to a 3% dose difference (in EUD) to the PTV, then it implies, from the EUD curves presented here, that the distance the PTV could move is up to 5.5 mm. This distance would be even larger for the CTV as the dose coverage is of the PTV, meaning that small movements of the CTV do not move it out of the high dose area. Thus a 3% acceptable dose difference would mean a large allowable movement. If this movement is larger than the CTV to PTV margin, then it implies that the PTV margin could be reduced. The current PTV margin ensures that the CTV stays within the 95% isodose when it moves. The EUD analysis here shows that even if part of a structure moves outside of the 95% isodose (in this case the PTV), the EUD does not fall rapidly. Therefore, if the PTV margin were made smaller, the CTV would occasionally move outside of the 95% isodose, but the EUD

would not fall as much as expected, and the size of the PTV margin could be made such that the fall in EUD to the CTV is never more than 3%. Reducing the PTV margin clearly has benefits to the patient, as the irradiated volume between the CTV and PTV boundaries is determined as healthy by the oncologist. Reduction of irradiation of this area will reduce normal tissue complications and improve patient outcomes.

The reasons for this are easy to realise. The dose distribution across a target volume is planned to be as uniform and flat as possible. Therefore moving around within the target volume should cause no change in EUD, which is a generalized average. The changes occur because the edges of the structure move into lower dose areas. However, the edge of the target volumes was defined here as the 95% isodose. The dose gradient either side of the 95% isodose is close to 1%/mm, and as the target structure moves, one end gets a dose higher than 95% and the other gets slightly lower. Thus for small movements while the dose profile is still rising to the 100% level, the EUD does not change much as there is a volume-averaging component to it. As the EUD parameter $a$ moves towards plus or minus infinity, the volume effect declines until the EUD is a measure of maximum or minimum dose respectively, but for values of $a$ within about -20 to +20, the volume effect is still present. The average dose across a structure is shown by the EUD when $a = 1$, and figure 5.4 shows that for movements up to about 8 or 9 mm, the average dose across the cylinder changes by no more than 1%.

The target volumes here were defined as conforming exactly to the 95% isodose, with the exception of the target used for the 3-field real dose distribution. In this case, the 95% isodose was roughly square with slightly rounded corners, and the target used was an ellipse fitted inside the square. This is a typical situation for target volumes such as prostates, where a 3 or 4 field plan is used. In this case, as shown by figure 5.8, the change in EUD is even less than for the targets where the 95% isodose conformed to the target edges. The reasons for this are also clear – there are areas of high, uniform dose

outside the target, so that when the target moves it encounters these areas and not the dose gradient down into the low dose areas. The effect is clearest for left and right shifts of the target volume in figure 5.8. The dose distribution from a parallel opposed pair of beams, as the left and right lateral beams are, is roughly uniform. As the target moves left or right, it encounters the edges of the anterior beam, but this makes up only 1/3 or so of the dose distribution. Thus the dose falloff to the right or left is shallow, and this combined with the oversized 95% isodose explains why the EUD changes so little for movements up to 8 mm – just 1%.

Figure 5.9 shows very different predictions for the change in EUD for organs at risk. Indeed, the change in EUD of the 'spine' structure of 21% for an anterior movement of just 5 mm is not a welcome sign to treatment planners. The change in EUDs for normal tissues is more difficult to study as the number of variables is much larger. For studying target volumes, the 95% isodose can be conformed to the target shape and a near-uniform dose profile is usually created within the target. For normal structures however, the size, orientation and location relative to the target will all alter the dose distribution they are in and the gradients in each direction. The value of the parameter $a$ also varies widely for normal structures, from 1.2 for organs with highly parallel functionality such as lungs (Burman *et al* 1991) to 20 for more serial organs such as the spinal cord. The two structures studied here were chosen to represent two clinically relevant IMRT sites, namely a rectum near a prostate and a spinal cord near an arbitrary tumour. In practice the spinal cord may be closer to the tumour, and the rectum typically overlaps with a prostate target once the prostate has had margins added to it.

The reasons behind the drastic changes in EUD are reasonably clear. Both structures studied here sat just outside the high dose region, but within the high dose

gradient areas. Unlike the target structure, which experienced a rise in dose at one edge and a fall at the other for small movements, the structures here experience an overall increase in dose. For the 'spine' structure, especially, a movement anterior in the dose distribution moves it into the penumbra of the two lateral beams and a rapid increase in dose is seen. The relative sizes of the structures also play a part. A small movement is comparatively larger for the rectum and spine structures than for the target. A 10 mm shift for the spine structure moves the entire structure into an area of dose that the original placement of the structure did not experience. For such small structures, steep dose gradients are more important as there is no volume averaging over areas of lower gradient while the structure is on the steep gradient. For a very large object, such as the target, the width of the high gradient area is small and there will always be a volume effect that takes in the lower gradient areas to either side.

The curves of EUD change in figure 5.9 clearly show that for critical organs at risk near to target volumes, assuming an equivalence of % dose to mm distance is not correct, and small movements may equate to large changes in equivalent dose. The conclusions that were drawn for target volumes are inverted for the case of organs at risk. If a 3 mm margin is allowed, the dose difference can be allowed as high as 10.8% for the 'spine' structure, or 6.8% for the 'rectum'. Here we only look at the increases in EUD, as unintentional decrease in dose to normal tissue is considered to be an unplanned bonus. Alternatively, if we set a dose difference tolerance of 3%, the allowed movement is as small as 1.2 mm. The problems are easier to see here – if dose verification is performed by visually comparing isodoses, and a 3 mm difference is allowed, the spinal cord structure in this case may receive 11% of the target dose extra. For a target structure receiving 65 Gy, this equates to a 7 Gy increase in dose. If the treatment plan already has the structure receiving close to its tolerance level of dose, this apparently small movement can have clear implications.

The work presented here relies on the use of a mathematical model of dose distribution as an indicator of radiobiological effects. Authors have shown (Wu *et al* 2002, Li *et al* 2003) that the generalized EUD formula is a good approximation to the more complicated full EUD equation first presented in 1997. Li *et al* also showed how the generalized EUD equation can in fact be derived from NTCP equations. Fitting the parameters of the NTCP equations to collected data from years of clinical radiotherapy has shown that the phenomenological results can be approximated well by the equations (Emami *et al* 1991, Burman *et al* 1991). However, despite the work relating simple equations to measured outcomes of radiotherapy, the EUD formula used here is a very simple one, and the absolute accuracy of the results it suggests should be taken with caution. The work presented here deals with the relative effects of moving a defined structure within a radiation dose distribution, and the changes it shows should be more accurate than the absolute doses calculated by the equation.

The results and analysis presented in this chapter show that there is room for improvement in the verification and quality assurance of IMRT. Improving the accuracy of IMRT is not simply about the way that the IMRT is delivered, but about how that delivery is verified – there is little to be gained in improving the consistency and accuracy of dose distributions if the checking methods still show that they are not correct. The concept that a dose difference does not change linearly with spatial movement should be clear to all those who carry out IMRT and treatment planning, as the dose distributions used in radiotherapy do not have constant gradients. It is therefore crucial to carry out work as detailed above in order to assess the way that dose and distance affect the outcome of radiotherapy.

The measure of EUD is a fairly simple tool to use for the prediction of radiobiological outcome. More work is clearly implied by the findings presented here.

Values of EUD can be related to values of NTCP and tumour control probability (TCP), and indeed, one needs an equivalent uniform dose across a structure before NTCP and TCP equations can be used. By relating dosimetric changes to values of control and complication probability, a much better understanding of changes in dose distributions can be obtained. However, the derivation of NTCP and TCP parameters and percentages is a much more site-specific and case-specific problem, whereas the work carried out here aims to present a more generalised introductory approach to the subject of relating dose distributions to changes in patient outcome. Thus further work clearly needs to be carried out by looking at specific tumour sites for specific patient cases if clinical outcomes of radiobiological response are to be integrated and implemented with physical measures of dosimetric evaluation.

**Conclusions**

Dose evaluation metrics such as the Kappa and Gamma indices combine dose difference and distance to agreement into a single evaluation criterion. The values for the DTA and dose difference are typically selected to be numerically the same, such as 3%/3mm or 4%/4mm. The work here shows, by using a biological indicator, that in the majority of practical structures modelled the effective dose to structures does not change by 1% for each millimetre movement of those structures relative to a dose distribution. For target volumes, the change in dose is smaller, suggesting that larger DTA tolerances could be used without compromising the patient outcome. For critical structures, the effect can be the opposite, suggesting that distance tolerances need to be much tighter to avoid potentially damaging irradiation of the organs.

This work shows that the common practice of setting equivalent distance and dose tolerances is invalid, and raises the question of whether purely physical dose evaluation, such as looking at dose and distance without considering the effect on the

patient, is valid or effective. In order to improve patient outcomes, it is important to move away from physical quantities such as dose or distance, and the next chapter will investigate using EUD as a radiobiological verification tool for IMRT. The possibilities of better, biologically-based dose evaluation will be discussed in depth in chapter 7.

# Chapter 6 – Modelling patient outcome as a function of systematic linac errors

## Introduction

In chapter 4, dosimetric errors as a function of systematic changes in linac behaviour were simulated. The dose differences induced by changes such as MLC positioning errors, poor calibration of MLCs and non-linear dose output of the machine were analysed from a purely physical point of view. Dose differences, distance to agreement and composite analyses such as the Kappa or Gamma indices only show how a dose distribution changes. Although the composite indices can show if two dose distributions are within certain combined tolerances of each other, all of the evaluation methods deal purely with the delivery of dose, and do not take account of how the changes in a distribution, whether a physical movement or a change in dose, affect the patient to whom the treatment is given. The aim of this work is to analyse, using simple radiobiological models, how the dose differences simulated and predicted in chapter 4 will affect patient outcome, and show how those patient outcomes might be improved through improving the accuracy (or not) of the physical dose that is delivered.

Chapter 5 introduced the concept of the Equivalent Uniform Dose, as first proposed by Niemierko in 1997. As with chapter 5, the work here will deal solely with Niemierko's revised concept of EUD, proposed in 1999 and based on the power-law response of biological systems. The EUD provides a single measure of the approximate dose that is delivered to a target or healthy structure when the dose across that structure is heterogeneous. It provides a way of assessing the overall effect on a structure of areas of dose inhomogeneity and hot and cold spots. Once a uniform dose for a structure is known, the approximate TCP or NTCP can be calculated, which then indicates the outcome effects for a patient.

The NTCP has been modelled by Burman *et al* (1991) as a sigmoidal function dependent on the uniform dose to a healthy organ. The sigmoid shape was fitted to the data collected by Emami *et al* (1991) to provide the parameters that define the shape of the sigmoid. The NTCP equation relies on uniform dose to an irradiated volume, and thus the dose variable to be inputted is the EUD. Thus from the EUD an NTCP percentage can be directly calculated. The data to which the original parameters were fitted by Burman *et al* are collected from many years of clinical experience and therefore represent realistic expectation of the complication probabilities.

The TCP is a more difficult concept to calculate an exact value for. The TCP depends critically on clonogenic densities, repair rates of cancerous cells and proliferation rates in the same way that the original formulation of EUD in 1997 did. However, estimates can be placed on the TCP, as it is assumed in any radical radiotherapy treatment that the prescribing oncologist is aiming for a reasonable chance of control.

The work presented here aims to deal with the physical dose differences shown in chapter 4 and equate them to patient outcomes, not only to show how those outcomes are affected, but also to determine if the apparent large dose delivery errors that fail tolerance levels actually matter when a patient is treated. By using calculated values of NTCP and TCP, the probability of uncomplicated control (PUC) as discussed in chapter 1, can be calculated. The PUC represents a good measure of the patient outcome, although it is not the only method of determining the efficacy of a radiotherapy treatment. The PUC describes the probability that the target tumour will be controlled without any complications of nearby healthy organs. In certain cases a level of normal-tissue complication may be tolerated if it provides for better dose delivery to the tumour and therefore better control. The therapeutic gain (TG) will also be calculated. This

expresses the difference between the NTCP and the TCP in a more direct fashion, which would enable an oncologist to see that there is an $x$% better chance of controlling the tumour than causing normal tissue damage. Both the PUC and the TG provide ways of evaluating patient outcome. Although the TCP is not being accurately calculated, both measures provide a useful tool for evaluating how heterogeneous dose distributions will change the endpoint of the treatment and how to improve it.

## Materials and methods

### Treatment plans

As with the work presented in chapter 4, this work will focus on two clinical IMRT treatment plans. The plans used are those used in chapter 4, namely a 50-segment 5-beam treatment plan for a prostate cancer patient (Patient H) and an 89-segment, 5-beam plan for a head-and-neck case (Patient S). Both treatment plans were segmented using the sliding window technique. Patient H's plan was created with Corvus and Patient S's with Xio.

### Error simulations

Various systematic errors were simulated with the treatments using the ICE software. For each of the two treatment plans studied, the following changes were made:

1 – Movement of the MLC leaves into and out of the field by various distances up to 2 mm for both Patient H and Patient S

2 – Simulation of the movement of MLC leaves under gravity by up to 3 mm for both Patient H and Patient S

3 – Movement of the MLC leaves under gravity with an added deliberate inward movement of the leaves of half the gravity movement for both patients

4 – Changing from a machine with perfect dose linearity to simulations of LA7 and LA8 at Addenbrooke's, for which the dose linearity was measured in chapter 4.

For each test carried out, the change in EUD from the unaltered state will be calculated, and the changes in tumour and tissue response probabilities evaluated.

**Dose differences**

To calculate dose distributions and differences, the independent dose calculation engine developed in chapter 4 is again used. The dose distributions calculated and analysed in that chapter are re-used for many of the analyses carried out for EUD analysis, and some new ones are calculated. More in-depth details of the dose engine can be found in chapter 4.

**EUD**

Once dose differences are calculated, they need to be converted into EUD. Niemierko's revised EUD (1999), also known as the generalized EUD is described on page 114 and is given by:

$$EUD = \left( \frac{1}{N} \sum_{i=N} D_i^a \right)^{1/a}$$

The EUD calculation function written for chapter 5 using the MATLAB programming environment is utilised for calculating the EUD of structures for which dose distributions have been calculated using ICE. The value of the EUD parameter $a$ used here is the same as that for the structures investigated in chapter 5. Patient H has the rectum near to the target as an organ at risk (OAR), and Patient S has the spinal cord near the target, both structures will be analysed for changes in EUD. The value used for

| | $TD_{50}$ | n | m |
|---|---|---|---|
| Spinal cord | 66.5 Gy | 0.05 | 0.175 |
| Rectum | 80 Gy | 0.12 | 0.15 |

**Table 6.I:** Parameters used in the NTCP calculations for the spinal cord and rectum. The data is taken directly from Burman *et al* (1991)

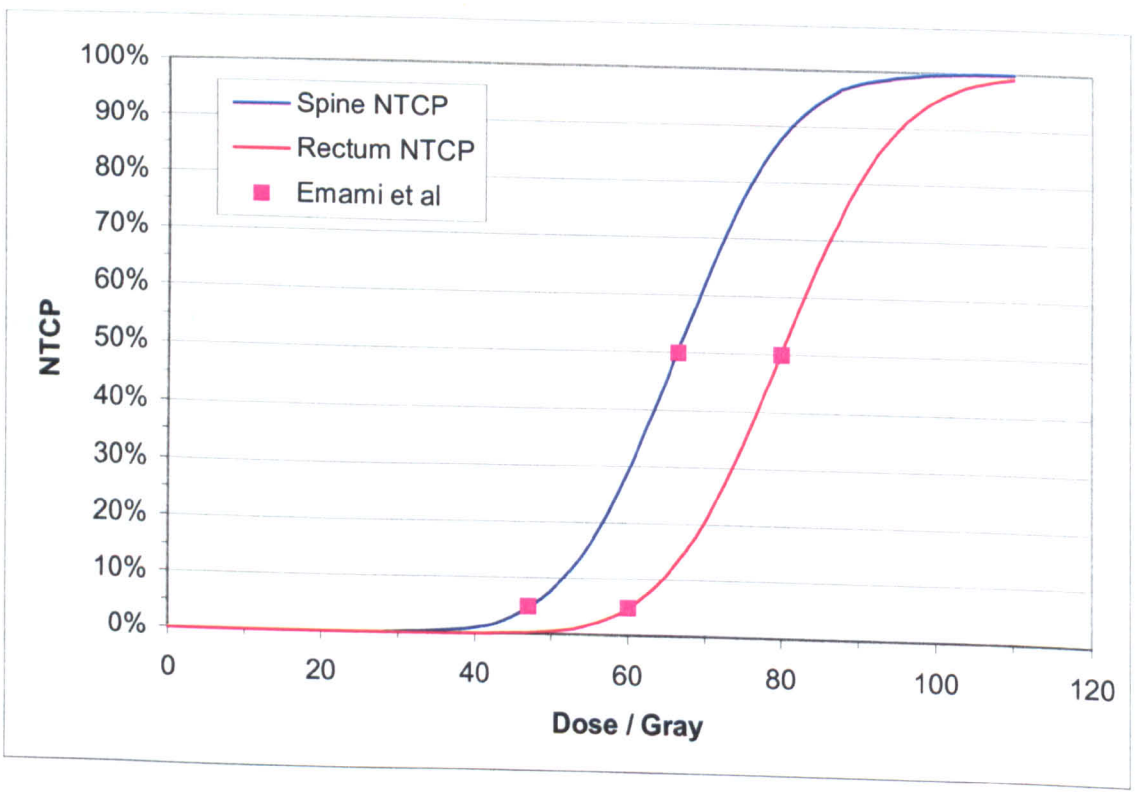



**Figure 6.1:** Graph showing the NTCP curve as expressed by Burman *et al* for the spine and the rectum. The pink data points show the data collected by Emami *et al*, the data to which Burman *et al* fitted the parameters of the curves.

the target is $a$ = -15, with $a$ = 8 and $a$ = 20 for the rectum and spine structures respectively.

### TCP, NTCP, PUC and TG

To assess likely patient outcome, the EUD will be used to evaluate the tumour control probability and the normal tissue complication probability. The NTCP will be calculated using the equation given by Burman *et al* (1991):

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-t^2/2}\, dt$$

$$t = \frac{D - TD_{50}(v)}{m \times TD_{50}(v)}$$

$$TD_{50}(v) = TD_{50}(1) \times v - n$$

$$v = V / V_{ref}$$

In this equation, $D$ represents the (uniform) dose to a fractional volume, $v$, of the reference volume $V_{ref}$. $TD_{50}$ is the dose that causes a 50% NTCP, m relates to the gradient of the NTCP curve and n is a volume-dependence factor. The parameters $n,m$ and $TD_{50}$ are taken from Emami *et al* (1991) and are shown in table 6.I. The volume-dependence factor $n$ is the inverse of the EUD parameter $a$, which determines how sensitive a structure is to hotspots. A small value of $n$, corresponding to a large value of $a$ implies that the structure is sensitive to maximum doses, regardless of the volume irradiated to that dose, whereas a large $n$/small $a$ implies that the structure is not sensitive to small volumes of dose change, only the dose the overall organ gets.

The NTCP is implemented using a pair of error functions in Excel[1], using one to model the curve below 50% and the other for NTCP greater than 50%. The NTCP equation shows that $t$ is equal to 0 at $D$=$TD_{50}$ and therefore a single error function can only model half of the NTCP curve because Excel does not allow a negative lower limit on the integration. The error function is symmetric so can be modelled as a pair of half-

[1] Microsoft Corporation, Redmond, WA

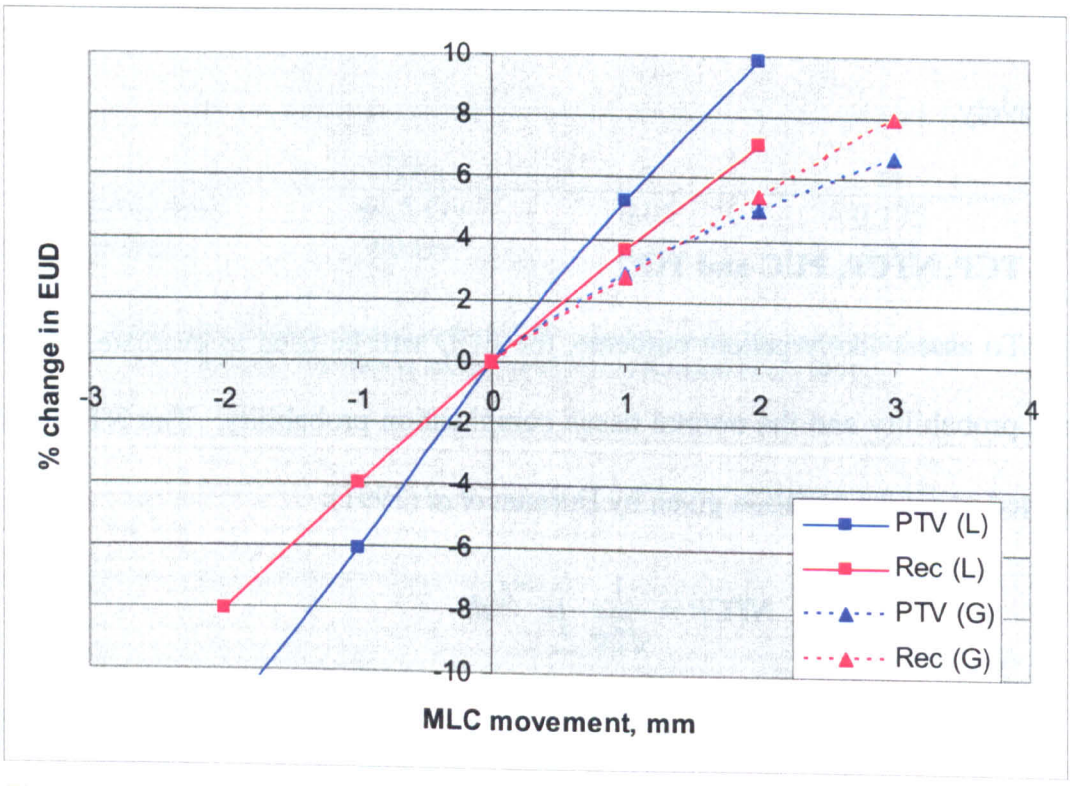



**Figure 6.2:** Graph showing the change in EUD for Patient H's PTV and rectum. The different lines represent MLC movements of both banks of leaves (L) or gravity motion of the lower leaf bank (G). The EUD change is a percentage of the EUD of the PTV for 0 movement

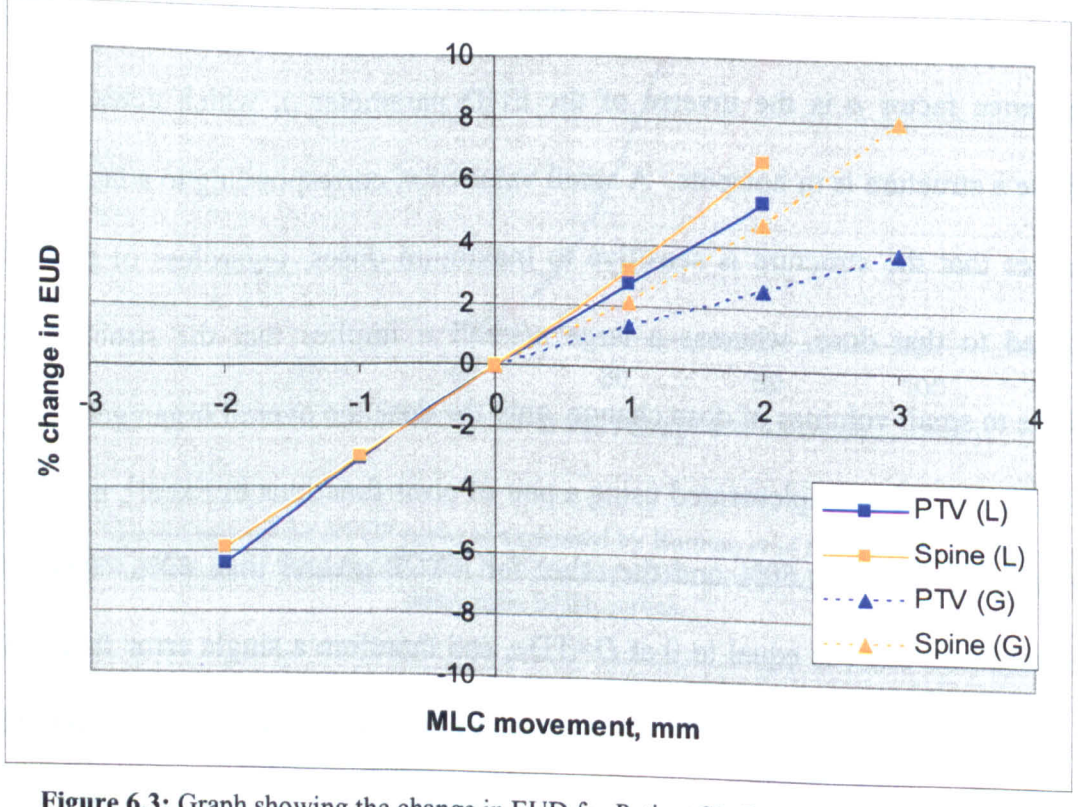



**Figure 6.3:** Graph showing the change in EUD for Patient S's PTV and spinal cord. The different lines represent MLC movements of both banks of leaves (L) or gravity motion of the lower leaf bank (G). The EUD change is a percentage of the EUD of the PTV for 0 movement

range functions. The parameters $n$, $m$ and $TD_{50}$ were inputted into the implemented equation and checked against the data of Emami *et al*, as shown in figure 6.1 (previous page). For all of the calculations of NTCP, the EUD is used as the input dose. The EUD is the uniform dose over the whole organ, and thus the volume factor used in the NTCP equation is always 1.

The PUC is given in chapter 1 as the 'healing factor' as expressed by Holthusen (1936), and is given by:

$$PUC = TCP(100\% - NTCP)$$

The therapeutic gain is calculated simply as the difference between the TCP and NTCP:

$$TG = TCP - NTCP$$

The TCP was not calculated precisely as was the NTCP, but instead approximated. The TCP is assumed to be 50% for the prescribed dose to the tumour. This places our point of interest in the centre of the sigmoidal tumour response curve, where the response to dose is approximately linear. The gradient of the TCP curve at this point, the $\gamma_{50}$, is assumed to be 2% per Gray and thus the TCP is assumed to change at this rate for changes in EUD.

**Results**

The dose differences and Kappa analyses for the various MLC movements simulated for these tests are presented in chapter 4. The EUD was calculated for the PTV and OAR for each patient, and the changes in EUD as the MLCs are moved are shown in figures 6.2 to 6.5. The changes in EUD are quoted as a percentage of the EUD of the PTV for the perfect plan. This normalisation is analogous to the way that dose differences are quoted, as a percentage of the prescription dose. It also means that

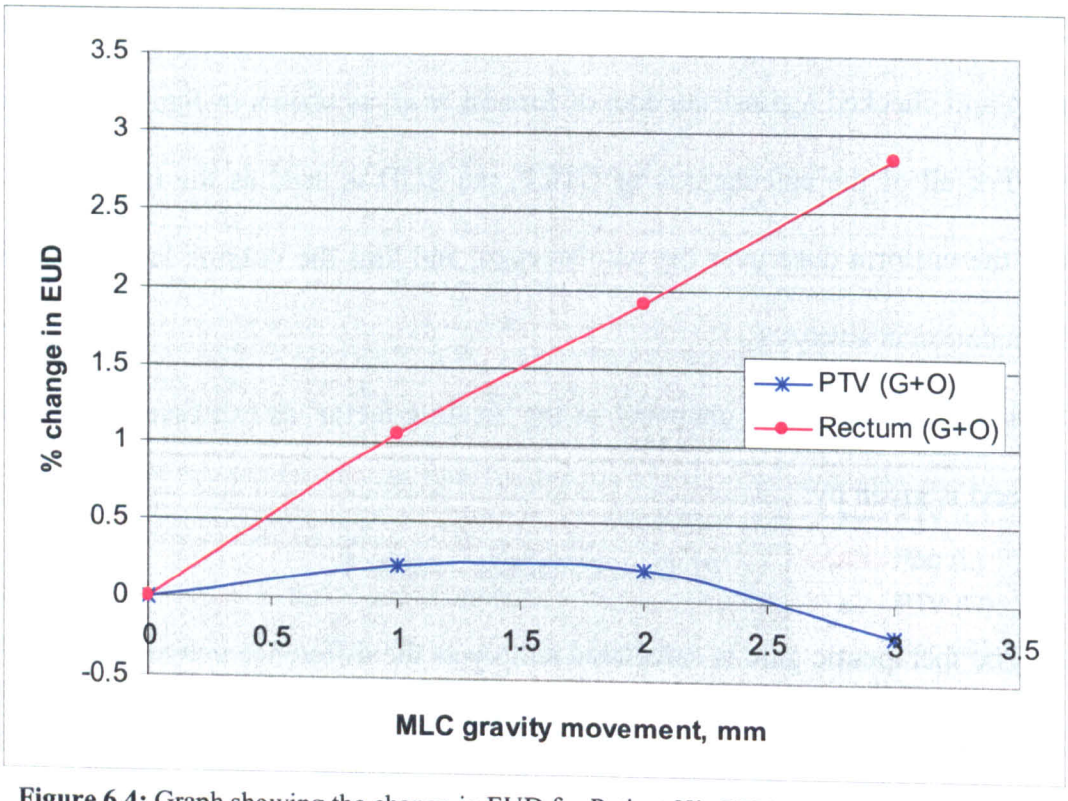



**Figure 6.4:** Graph showing the change in EUD for Patient H's PTV and rectum for a gravity movement shown with an MLC offset of half that amount on both banks of leaves.

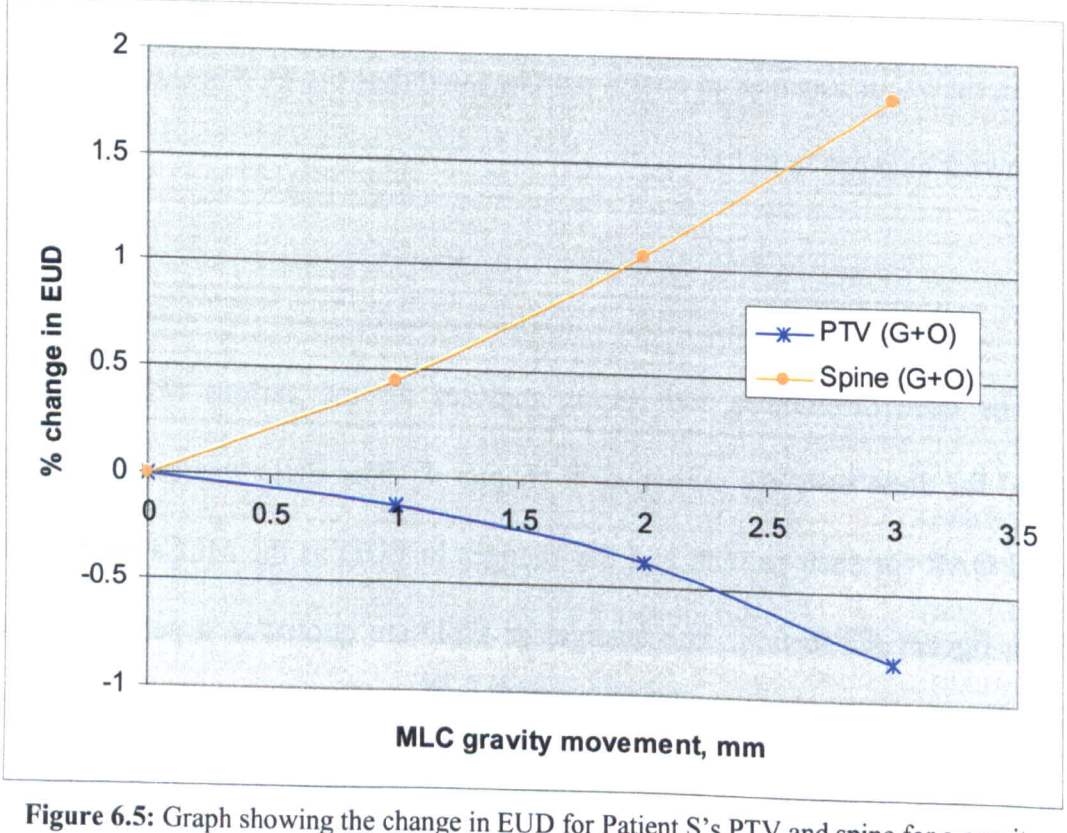



**Figure 6.5:** Graph showing the change in EUD for Patient S's PTV and spine for a gravity movement shown with an MLC offset of half that amount on both banks of leaves.

the changes are invariant to the actual dose planned, even for the organs at risk. For Patient H the prescription dose was 60 Gy, with an EUD to the PTV of 61.5 Gy; Patient S had a prescription dose of 65 Gy with an EUD to the PTV of 64.6 Gy.

Figure 6.2 (previous page) shows the change in EUD for Patient H's tumour and rectum as the MLC leaves are moved away from perfect alignment. The graph shows changes for when both banks of MLCs are moved out of the field and for when the leaves move under gravity during gantry rotation. Negative movements of the MLCs indicate that the leaves are moved into the field. The graph clearly shows that as the leaves move out of the field, the EUD of both the PTV and rectum increase, and the EUDs decrease if the leaves are moved into the field. This is to be expected, as larger fields mean that more radiation is delivered, and the average dose will increase regardless of how that average is calculated.

Figure 6.3 (previous page) shows the same data, but for Patient S's treatment plan, again showing that the EUDs of the spinal cord and PTV increase as the fields get larger.

As with the work in chapter 4, deliberate offsets were added to the MLC leaves to partly counteract the gravity motion. The leaves were set to be too small by one half of the simulated gravity motion, and the changes in EUD were calculated from the dose distributions generated. Figures 6.4 and 6.5 show the changes in EUD for Patient H and Patient S respectively. The horizontal axis shows the amount of gravity movement, and the deliberate MLC offset is half of that value at any point on the graphs. Both graphs show that the changes in EUD have been reduced significantly compared to figures 6.2 and 6.3 for both the PTVs and the organs at risk.

From the evaluated EUDs, NTCP values were calculated for the organs at risk. TCP was also estimated based on a 50% value for the perfect plan EUD and $\gamma_{50}$ of 2% per Gray. As the MLC leaves were allowed to move under gravity or were moved in or

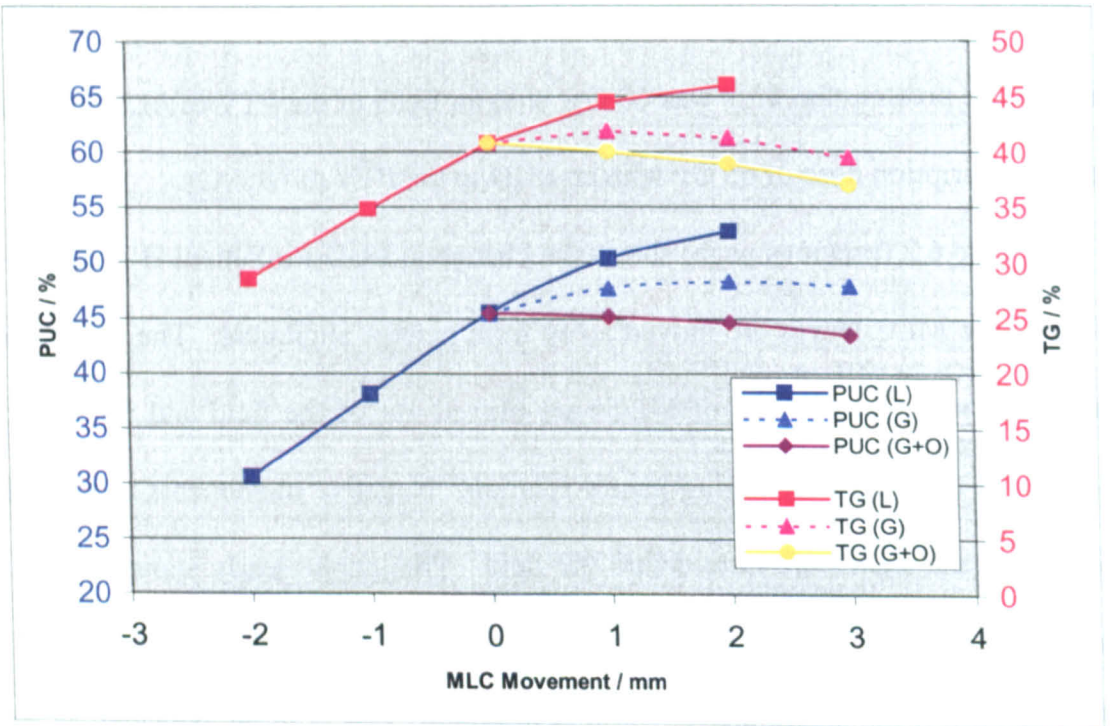



**Figure 6.6:** Patient H's outcome as a function of MLC movement. The graph shows the probability of uncomplicated control (PUC) and Therapeutic Gain (TG) for plan simulations in which both banks of leaves are moved or one bank moves under gravity (L and G respectively). Also shown are the PUC and TG for the situation of gravity movement with a deliberate MLC offset (G+O).
Positive MLC movements indicate movements out of the radiation field.

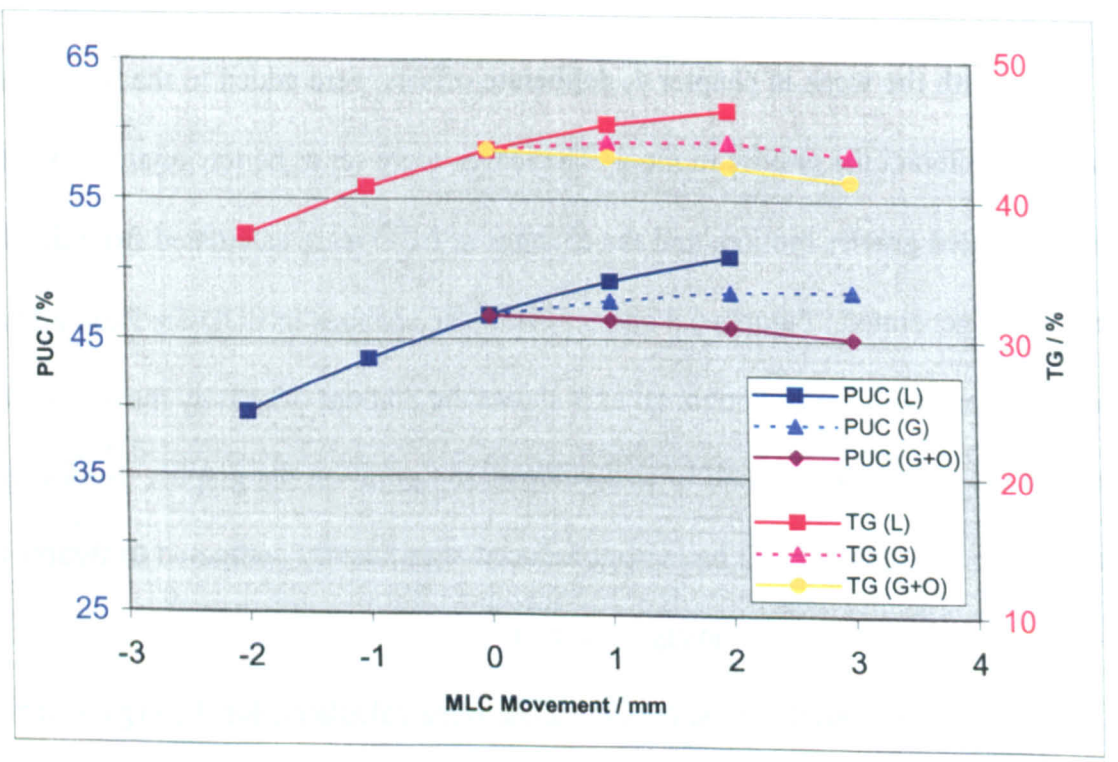



**Figure 6.7:** Patient S's outcome as a function of MLC movement. The graph shows the probability of uncomplicated control (PUC) and Therapeutic Gain (TG) for plan simulations in which both banks of leaves are moved or one bank moves under gravity (L and G respectively). G+O represents the case of gravity-induced motion with a deliberate offset of the MLC leaves.
Positive MLC movements indicate movements out of the radiation field.

out of the field, the EUD for both the PTV and organs at risk changed, as shown in figures 6.2 and 6.3. The PUC and therapeutic gain were both calculated and are shown in figure 6.6 for Patient H and figure 6.7 for Patient S. Each graph shows the change in PUC and TG for three situations: both MLC banks moving out of or into the field (positive and negative x-axis); MLC leaves moving under gravity; and MLC leaves moving under gravity with a half-distance offset into the field.

Finally, figure 6.8 (next page) shows the effect on PUC and TG as the dose linearity of the treatment machine is varied. The dose linearity for machines LA7 and LA8 at Addenbrooke's, as measured for chapter 4, was applied to the dose calculations and the PUC and TG were calculated from the different plans. Figure 6.8 presents the PUC and the TG for both Patient H's and Patient S's plans, and shows small but measurable changes.

**Discussion**

Chapter 4 showed that small changes in MLC positioning accuracy could lead to small areas of large dose deviation from that expected from a machine with perfect MLCs. The analysis methods showed that areas of the target volume could easily fall outside of a 3%/3 mm tolerance with only small changes in MLC accuracy. The work presented in chapter 4 deals solely with the physical aspects of dose in terms of the locations of dose and the exact values of it, rather than the biological effects. It is entirely plausible that a small area of high over or under dose could have negligible impact upon patient outcomes or that a large area of low over or under dose could have a large impact.

Figures 4.12 and 4.14 showed that the percentage of the PTV that fails the Kappa index (3%/3 mm) test is not linear with MLC movement and exhibits a very steep increase as the amount of movement is increased beyond 1 mm. However, such a

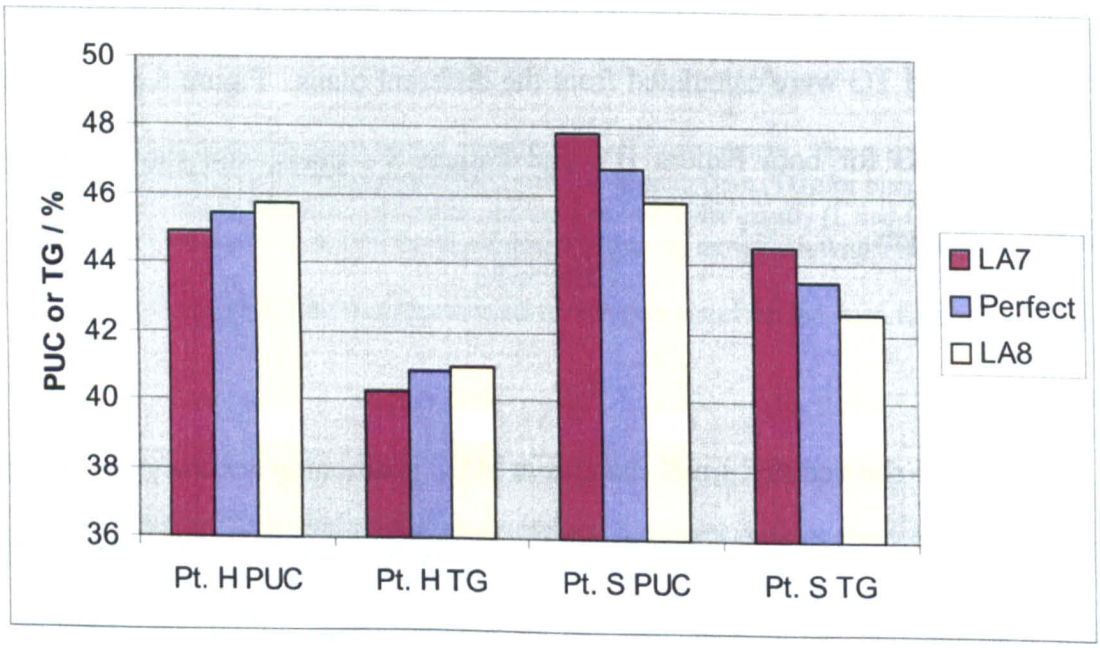



**Figure 6.8:** Changes in PUC and TG for patients H and S when treated on different treatment machines which have different dose linearity: A perfect machine and two machines at Addenbrooke's, LA7 and LA8.

graph does not give any information as to the nature of that failure, whether the failing areas have very high values of Kappa or values that are only just larger than 1. The Kappa histograms provide more information, showing a tail of Kappa index values that increase up to $\kappa$ =2.5 for a 2 mm movement of the MLCs for Patient S's plan. The top of the Kappa histogram also highlights the potential issues as the histogram does not begin to fall until a Kappa index of nearly 0.9, showing that the whole PTV area has a Kappa index of at least this value. Still, this does not provide information on how the dose differences affect the patient because Kappa is sign-independent, in that the histogram does not show if doses are too low or too high.

Figures 6.2 and 6.3 help to address this issue. Figure 6.2 shows that for an MLC movement of 2 mm, the change in EUD for Patient H's PTV is around 10%. This shows that the tail of Kappa index values shown in figure 4.11 that extend to $\kappa$ =2, increase the EUD by more than the index suggests. A Kappa index of 2 implies a 6% dose difference or a 6 mm distance to agreement. Figure 4.11 shows that this is the maximum Kappa index value associated with a 2 mm MLC offset, and so the EUD might be expected to change by no more than 6%. However, the composite indices such as Kappa and Gamma allow large differences in dose if the distance to agreement is small, in high dose gradient areas. Although these distances may be taken into account in planning margins in terms of ensuring coverage of a target, the dosimetric implications are larger than the composite indices suggest. A tumour, or healthy tissue, does not know that the correct dose is being delivered a short distance away, and the only effects that an area of tissue feels are those caused by the local dose. Thus if a structure has a dose gradient across it, a change in dose may register a small DTA, but it is still a change in dose. The EUD analysis shows that the biological outcome of the changes in dose distributions is not necessarily as small as composite indices suggest. Indeed, the Kappa index as used in chapter 4 is less tolerant of composite dose and

distance errors (as explained in chapter 3) and is thus reporting higher values of the index than would the Gamma index method. Therefore the changes in EUD seen here are even larger when compared to the results of analysis by the Gamma index than when compared to the Kappa index.

Figure 6.2 shows that for a 1 mm MLC offset out of the field, the EUD for the PTV increases by 5.3% and that of the rectum increase by 3.7%. Since the EUD is the equivalent dose that is delivered to the *whole* volume, then it can be equated to a 3.7% dose difference everywhere across the structure. If, as is the case for target volumes, the dose gradient is low, the DTA will be high and the whole volume will have a Kappa index greater than 1 and hence the whole volume will fail the Kappa test at 3% tolerance. From figure 6.2 it can be seen that 0.55 mm is the maximum MLC movement that can be made to both leaf banks before the EUD of the PTV changes by 3%. This is a much smaller movement than that suggested by the work in chapter 4, which showed that 0.6 mm of movement would cause only 1% of the PTV to have a Kappa index greater than 1 with a 3% tolerance and that even 2 mm of MLC movement would only cause 75% of the PTV to fail Kappa.

This is because of the way that EUD averages the dose over a volume. Chapter 4 shows that a small proportion of the target has a Kappa index of greater than 1, but that this proportion has a dose that is very much more than that planned. Additionally, there are many areas that have a Kappa index of less than one but which have a high dose, as seen by the rapid increase in failure for slightly larger MLC movements. The EUD averages all of these hotspots and shows that they equate to an equivalent raising of the dose to the whole volume, which in this case is by more than 3%. Interestingly, figure 6.3 shows that for Patient S, 1.1 mm is the maximum MLC offset that can be made before the EUD changes by 3%. This implies that the sensitivity of Patient S's plan is lower than that for Patient H, which is contrary to the results found in chapter 4.

Where that showed that a greater percentage of Patient S's PTV failed the Kappa 3/3 test for the same MLC shift, this EUD analysis shows that the effect on the patient is smaller.

For Patient H, the change in EUD for the PTV is roughly halved between MLC offsets to both banks of leaves and gravity movement, which will affect only one bank of leaves at a time for most gantry angles and for all gantry angles used in these plans. Interestingly, the change in EUD for the rectum is not as different between gravity movement and MLC offset. The reduction in EUD change between MLC offset and gravity movement is around 25% rather than 50% as it is for the PTV. This is explained by the location of the rectum itself, which is posterior to the PTV. Since gravity movement of the MLC leaves affects only the lower bank of leaves, the posterior areas of a patient (when lying supine) will receive more changes in dose than the anterior sections. This is discussed more in depth in chapter 4, but is caused because some of the largest changes in fluence for intensity modulated beams are not in the middle of the beams, but at the edges, where the intensity rises from nothing outside of the irradiated area to some value inside. Thus the largest dose errors are seen at the beam edges, and at the lower beam edges for gravity movement. The rectum will not be affected very much by changes to the MLCs on the tops of the fields, and this explains why the EUD change is nearly the same for MLC offset and gravity movement. A similar effect is seen in figure 6.3 for Patient S, whose OAR, the spine, is also posterior to the PTV and is thus more affected by the lower banks of MLC leaves than the top bank. As with Patient H, the change in EUD for the PTV is approximately halved when only considering gravity motion rather than an offset to both MLC banks whereas the change in EUD for the OAR is only reduced by 25% when considering gravity motion over both banks moving.

Figures 6.2 and 6.3 show that the changes in EUD are approximately linear with MLC movement. This reflects the fact that the EUD formula is a generalized average and that opening an MLC leaf by 2 mm causes double the dose difference as opening by 1 mm. The EUD also takes into account maximum and minimum doses for OARs and targets respectively, but the maximum change in dose is also linear with small MLC movement which exposes more of the linear section of the beam penumbra.

Figures 6.4 and 6.5 show the effects that introducing a deliberate MLC offset has in countering the gravity movement. Since the gravity movement is an inherent feature of the construction of the linac and the MLC leaves, it cannot be removed or reduced. By introducing an offset to the MLC leaf positions, it can be partially countered. Figure 6.4 shows that for Patient H, the offset removes virtually all of the EUD change for gravity movements as large as 3 mm, with no change larger than 0.25% seen for any movement up to this value. The change in EUD for the rectum is larger, at approximately 1% per millimetre of gravity motion. Compared to figure 6.2 though, the changes in EUD are much reduced from the 2.5% per millimetre associated with gravity movement alone. Although the change in EUD for the rectum is still positive, and will therefore lead to increased probability of adverse side-effects for the organ, the graph shows that for realistic values of gravity motion below 2 mm, the change in EUD is less than 2%, which should be considered as an acceptable limit, since it is below the percentage tolerance that is ascribed to much checking, including composite indices. In the light of the fact that composite indices such as Kappa and Gamma tend to allow dose differences that cause large changes in EUD (as discussed above), a 2% change in EUD corresponds to a much smaller composite tolerance, possibly as low as 1%/1 mm.

Figure 6.5 shows that applying an MLC offset to the gravity motion also reduces the EUD changes for Patient S's organs. As with Patient H's analysis, the changes in

EUD to the PTV are much reduced, below 1% for gravity movements as high as 3 mm. Again, for a realistic range of gravity motion up to 1.5 mm, the change in EUD to the PTV is below 0.4%. The change in EUD to the spine is also much reduced, from the 2% per millimetre shown in figure 6.3 to less than 0.5% per millimetre.

When the changes in EUD are converted into patient outcomes, there are similarities between the results and some surprising answers. Figure 6.6, for example, shows that for Patient H the difference in change of PUC between gravity-induced MLC movement and movement of both banks is approximately double, as with the EUD. The changes in PUC can also be virtually eliminated by applying a half-distance offset to the MLC leaves to counteract the gravity movement. What is surprising, however, is that the maximum PUC is not achieved when the MLC leaves are perfectly aligned to the plan, but when they are all moved out of the field. This may seem counter-intuitive, as the delivered plan is not accurate, and does not match the desired dose distribution. However, the reason for the improvement in patient outcome lies in the actual values of TCP and NTCP. The $TD_{50}$ for the rectum is 80 Gray, which is well above the prescription dose of 60 Gray for this plan. Although parts of the rectum overlap the PTV, no part of the rectum will receive more than 107% of the prescription dose, as required by ICRU report 62. Therefore the maximum dose that the rectum will receive is 64.2 Gray, although this will increase as the MLCs are moved from the ideal position. The NTCP for the rectum is therefore around 10%, rising to 20% for a 2 mm MLC movement. The TCP is assumed to be 50%. Therefore, if both the TCP and the NTCP are raised by 10%, for example if both are in their linear response ranges and the dose to both the PTV and the rectum rise by the same amount, then the new values may be 20% and 60% for the NTCP and TCP respectively. The PUC is calculated not by the NTCP, but by the inverse: 100%-NTCP. Therefore the value of interest to the PUC equation

has changed from 90% to 80% with a 10% increase in NTCP. The change from 90 to 80 is a smaller proportion than from 50 to 60 for the TCP, and thus the PUC actually increases, despite the NTCP doubling. Figure 6.6 shows that the TG increase as the MLC leaves are moved out the field, which actually shows that the TCP increases more rapidly than the NTCP, and thus the PUC will rise even faster than discussed above. The reason for the TCP increasing more rapidly than the NTCP is partly because the PTV and rectum structures overlap. The maximum dose to the rectum is likely to be similar to the maximum dose to the PTV, and the EUD of the rectum is largely controlled by the maximum dose. Thus an increase in dose to the PTV, as given by larger field sizes, will increase the EUD of the rectum and the PTV. The increase in dose will help to fill in areas of low dose in the PTV that were dragging the EUD down, and may well then increase the EUD of the PTV by more than that of the rectum, especially since some of the rectum is outside of the target volume.

Figure 6.7 shows similar results for the PUC and TG analysis of Patient S. The effect of gravity-induced movement is approximately half of that for both MLC banks moving, and the effects can be reduced significantly with a deliberate MLC offset. Also similar to Patient H, Patient S shows an increase in PUC and TG when the MLC leaves are moved away from perfect alignment and out of the field. Again, this is because the NTCP starts at a low percentage level, 6% in this case, because the dose to the spinal cord is nearly 20 Gray below the $TD_{50}$ of 66.5 Gray, although this is still a relatively high clinical NTCP for the spinal cord.

The doses involved for Patient H's treatment are relatively low compared to current practices, in which it is common to irradiate the prostate to 74 Gray. To investigate the approximate effects of changes in dose at this higher irradiation level, Patient H's treatment was scaled up by 23% to reach a tumour dose of 74 Gy, thereby also increasing the dose to the surrounding tissues. The NTCP and TCP values were

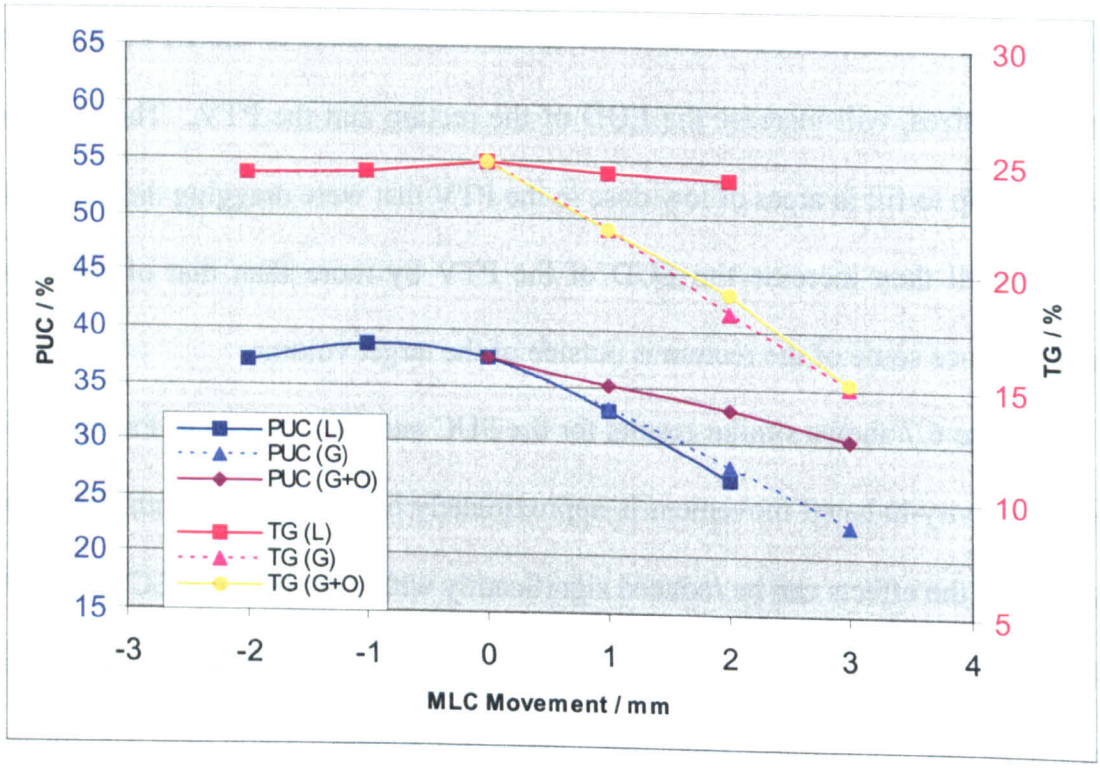



**Figure 6.9:** Probability of Uncomplicated Control and Therapeutic Gain for Patient H's plan when all doses are scaled by 23% to deliver 74 Gy to the tumour.

then recalculated. The NTCP was calculated using the previous formulae, while the TCP was assumed to have risen to 78 % (14 Gray increase at 2% per Gray). This is not an accurate picture of a 74 Gy prostate treatment as inverse planning systems will devise better ways of delivering the extra dose to the tumour while keeping the dose to the surrounding structures as low as possible. However, the differences between the 60 Gy and 74 Gy cases are still informative. Figure 6.9 shows the PUC and TG for the scaled-up dose case and shows some interesting differences when compared to figure 6.6. The first point to note is that the therapeutic gain dropped from 40% to 25%. Although the TCP has increased to an assumed 78%, the NTCP has increased from 9% to around 50% because the rectum EUD is now close to the $TD_{50}$ dose of 80 Gy. The effect on the PUC is similar, reducing from 45% to 37%, again mainly because of the much higher rectal doses incurred with extra dose to the prostate. These reductions in PUC and TG may be balanced by an oncologist against the fact that the TCP will be much higher at 74 Gy than 60 Gy and that the complications may be tolerable.

The most interesting change between figures 6.6 and 6.9 is the shape and direction of the curves as the MLC movement is changed. In figure 6.6 the NTCP was numerically low enough that any extra dose improved PUC and TG. For the 74 Gy case, the situation is much changed, and it appears that the current dose distribution is almost ideal. Any change in the MLC positions – both banks moving in or out, gravity movement or gravity movement with an offset – lead to a decrease in the therapeutic gain. In this case, the gravity movement leads to the biggest changes, as the NTCP of the rectum is most rapidly near the $TD_{50}$ dose. Adding an offset to the MLC leaves of half the gravity movement helps very little as the maximum dose to the rectum still increases and there is no associated increase in target dose. Curiously, moving all MLC leaves in or out of the fields changes the TG very little, as the TCP and NTCP increase or decrease at approximately the same rate, both being in the linear portion of their

response curves. The PUC shows that only very small improvements in calculated patient outcome can be achieved by moving the MLC leaves, with any enlargement of treatment fields leading to a decrease in PUC. Offsetting the gravity movement of the leaves does help to improve the PUC, although it is always lower than for the perfect MLC case. This however may be irrelevant in a clinical department as normal tissue toxicity is balanced against tumour control. In such situations a decrease in both NTCP and TCP is bad, as a drop in tumour control comes at the expense of a lowering of NTCP that was already acceptable.

The data in figure 6.9 suggests that, as expected, systematic changes in MLC positions can lead to deterioration of patient outcome and that improvements to the accuracy of IMRT delivery can clearly lead to improvements in outcome.

Figure 6.8 shows that the changes in PUC and TG when non-linear treatment machines are used are not always negative. In fact the data suggests that one machine is better for treating a patient, both in terms of TG and PUC, but that the ideal machine is different for Patient H and Patient S. Although the differences in PUC and TG are small, at no more than 2%, this graph again highlights that perfect accuracy need not imply the best outcome for a patient.

EUD, although based on radiobiological modelling, is a measure of a physical dose distribution, and tells us only the equivalency of a dose distribution, rather than what that dose distribution means for a patient in whom it is created. Thus EUD is a good tool for showing us how accurate a treatment delivery is compared to that planned in a way that does allow hot and cold areas, which are most likely inevitable in complex treatments such as IMRT. The EUD is thus a good tool for analysing if a dose distribution is close enough to that prescribed, and accuracy tolerances can be set on

measurements of EUD to determine if a plan is valid for treatment. However, a more in-depth radiobiological inspection of the data implies that improving accuracy and achieving exactly what the treatment plan requested does not always mean that the outcome of the patient is improved. Indeed, errors in the dose distribution have been shown here to inadvertently increase the probability of a successful patient outcome, although in some cases, the outcome is degraded. Thus the PUC and TG provide us with useful tools for assessing the outcome of treatments and can inform us on how to improve them.

The results presented here rely on data that is not accurately calculated, namely the estimation of the TCP at 50%. The exact value of TCP is difficult to calculate, as a radiobiological inspection of the mechanisms leading to tumour death is deeply involved and beyond the scope of this thesis. A prescribing oncologist will most likely prescribe a target dose that will give an expected TCP of greater than 50% for many cancers, especially if the known risks of NTCP are low. The analysis of results and data presented here remains valid so long as the TCP does reach very high levels. For example, the increase in PUC with increasing dose is largely due to the fact that a 10% increase is a larger proportion of the TCP (50%) than it is of 100-NTCP (~90%). If the TCP were larger then a 10% increase becomes a smaller proportion. So long as the TCP is not of the same size as the NTCP then the increase in PUC would still be seen. The tools and discussion presented here are therefore useful in showing that physical measures of dose distributions are not always the absolute measure that many have taken them to be. Much more in-depth analysis of dose distributions need to be undertaken in order to assess whether or not a difference between that delivered and that planned is actually harmful to the patient or if, as may well be the case, the difference actually improves the outcome for the patient.

## Conclusions

The work presented in this chapter has shown that radiobiological tools can be used in assessing both the accuracy and outcome of inhomogenous dose distributions and dose distributions that differ from that planned when they are delivered. The EUD is a useful measure of the accuracy of a plan, because it allows for hot and cold areas if they counteract each other and lead to the same equivalent dose delivered to the target tumour and organs at risk. The NTCP and PUC are useful tools in analysing how the outcome of a treatment depends on the dose distribution.

In terms of dosimetric accuracy, using the EUD tool reveals that a movement of the MLC leaves by just 0.6 mm can lead to a change in dose of 3% to the PTV, whereas more conventional analysis using the Kappa or Gamma index would suggest that only very small areas of the PTV would be out of tolerance (for 3% and 3 mm) with MLC movements even as large as 1 mm. This work shows that the EUD can highlight inaccuracy in dose distributions and therefore lead to improvements in that accuracy.

In terms of patient-specific outcome, the probability of uncomplicated control shows that there is the possibility of a better outcome for a patient when the dosimetric accuracy is not perfect, and that perfect dosimetric accuracy does not necessarily imply the best patient outcome. However, in cases where nearby organs at risk are receiving doses close to tolerance levels, perfect dosimetric accuracy is likely to give the best patient outcome. The effects are clearly patient-specific and depend on exact values of TCP and NTCP.

Further work is implied in studying more patient cases to analyse how accuracy and outcome are linked for the general patient population. From that, an idea of how accurate treatments must be to produce acceptable patient outcomes may be arrived at. Alternatively, the variability of data may imply that dosimetric analysis of every patient may be necessary if accurate information on outcomes is required.

# Chapter 7 – Discussion

IMRT is a complex process which involves many separate procedures, each of which able to introduce its own inaccuracies that can affect patient outcome at the end of the treatment process. Careful scrutiny is required to ensure that the best possible accuracy and the best possible patient outcomes are achieved, as many of the factors that affect IMRT are subtle and not necessarily intuitive. It can be the case, for example, that improved accuracy does not lead to improved outcome.

One of the most important aspects of radiotherapy is delivering the radiation beams to the correct part of the patient in order to build up the desired dose distribution. The work in chapter 2 shows how a non-invasive and non-ionising technique can be used to assess the accuracy of immobilisation devices such that treatment planning margins may be applied in a more informed manner. Previous studies have focused on retrospective measurement of patients on treatment. While this provides data on how real patients are being set up and will show how factors such as weight loss and patient compliance can make margins larger or smaller, the data may not be as complete or as useful as expected. Techniques for immobilisation are changing all the time and applying yesterday's margins to today's patients may not necessarily be useful in a rapidly changing environment. Additionally, published data may not be applicable to other centres due to differences in mould-room staff competency, radiography staff and exact fixation methods. Radiotherapy centres therefore need to carry out their own setup studies. The process described here allows that study to be carried out for several immobilisation devices in a safe and controlled way. The use of a single subject may not give absolute answers to the question of setup accuracy, but can provide excellent relative data, showing a radiotherapy centre which immobilisation device is best for the methods and procedures in use. In addition, the movement data acquired from such a

study can be used to guide (although not set) treatment planning margins, especially if the results are significantly different from those in use already. Once a particular immobilisation device has been selected as the best, the 3D digitiser method can be used in a more extensive study to acquire more accurate absolute movement data to guide the decision over treatment margins more objectively.

The method used in this study for measuring patient position deviation also differs to that used in many previous studies. This study investigated the changes in subject position relative to the treatment couch, which is clearly important when patients are set up using couch positions only. Previous studies have either not stated how the patient setup is carried out, or have aligned patients using skin or shell markers. This highlights the need for a centre to carry out its own study based on its setup procedures.

Patient setup is clearly important for all radiotherapy, but especially so for IMRT, where dose distributions tend to be far more complex, conforming not just to target volumes, but more critically to organs at risk too. A movement of the patient will move the target volume out of the high dose gradient area, negatively affecting tumour control. At the same time, the movement may move critical organs into higher dose areas, particularly if they are being spared radiation by nearby dose sculpting and high dose gradients. Therefore small, unaccounted for deviations in patient position can eliminate the advantages of IMRT and even make the treatment more dangerous than simpler conformal radiotherapy. The work presented here then is an important first step in improving the accuracy of IMRT delivery and therefore the patient outcomes as a result by identifying the best method of positioning patients to ensure that the radiation is delivered accurately.

IMRT verification depends on the use of software tools to analyse and present the data available to those carrying out the work. Chapter 3 showed that the most common dose evaluation metric in use, the Gamma index, has several different interpretations, which give different answers. The study also showed that the Gamma index is over-tolerant of dose and distance shifts in areas around the edges of tumours and where dose gradients are moderate. The introduction of the Kappa index brings a new quantitative evaluation tool that is less tolerant in these areas and the same as the Gamma index for very steep and very shallow dose gradients. The Gamma index is up to 41% more tolerant in moderate dose gradients, such that dose errors or distance errors of up to 1.41 times the desired tolerance are accepted as within tolerance. There is the possibility that the Kappa index will highlight extra 'problem' areas, and it may be the case that these areas are considered acceptable by subsequent human checks. However, from a patient-centric point of view, highlighting *potential* problems is better than blindly allowing them with a dose evaluation tool that is more tolerant than it seems by up to 41%. Despite this, the Gamma index is a widely-used tool, and so proposals for better calculation methods – both faster and more accurate – have been presented. The aim of radiotherapy is clearly to be as accurate as possible, but making the process easier and faster is also important. IMRT verification can be a long process, and so improving calculation speeds can help to increase the amount of work that staff are willing to do for each plan, or decrease the time taken, making it possible to treat more patients with IMRT.

Over the last 10 years, as IMRT has become more mainstream and accessible, much has been discussed about the need for accuracy through the IMRT process. Although it is intuitive that accurate treatment machinery is required, very little work has been carried out to show what sort of accuracy is required. Previous studies have

focussed on dynamic-MLC IMRT, as moving leaves clearly need to be near-perfect. There seems to be a belief that step-and-shoot IMRT is a simpler process and that therefore the leaf tolerances don't need to be so tight. Centres may choose step-and-shoot IMRT over dynamic if they don't believe that their MLC calibration is as good as it needs to be. However the work presented in chapter 4 shows that step-and-shoot IMRT is susceptible to large changes in dose delivery caused by small changes in MLC and dose output behaviour. Such systematic changes are typically overlooked in radiotherapy as there is a general belief that treatment margins cover many inaccuracies. With conformal radiotherapy, as discussed in chapter 3, 2 mm field errors simply mean that 2 mm of extra margin around tumours and organs at risk is required to ensure proper dose coverage. With IMRT, the situation is different due to the building up of fields that do not cover the whole target volume and the rapid changes in fluence across a beam's cross-section, where each change in fluence effectively becomes a new field edge. As an example, Siemens quote that the positional accuracy of their MLC leaves is ±2 mm for the older Primus linacs, and ±1 mm for the newer Oncor machines. Although this level of accuracy seems excellent, the dose calculations in chapter 4 show that a 1 mm deviation in MLC position, if it occurs to all leaves, can lead to tolerance failure of significant portions of the PTV and surrounding tissues.

Chapter 4 shows that the level of accuracy needed in MLC calibration is higher than may be expected in order to achieve the desired dose distributions. MLC leaves need to be calibrated to better than 1 mm, especially if highly complex treatments such as head-and-neck cases are performed. Measurement of systematic errors, such as gravity motion of MLC leaves and dose output changes of the linac are also important, as these can lead to significant changes in delivered dose. An accurate method of performing MLC measurement has been presented, which uses equipment that is typically already found in many radiotherapy centres. Using in-house software such as

that described, the process of using film to measure MLC positions can be fast and accurate, with precision better than that offered by electronic systems for static MLC measurements. Once a process such as this has been implemented, removal of calibration errors becomes simple. With sufficient accuracy of calibration, such as that possible with film-based measurements, work can be done to remove other systematic changes. Gravity-induced motion of MLCs cannot be removed without replacement of the MLC mechanisms, and so a linac is likely to have the same mechanical behaviour for the length of its service (typically 10 years in the UK). The work in chapter 4 showed that reducing the effects of the gravity motion can result in much better agreement between the dose delivered and that planned and therefore taking steps to counter such consistent systematic MLC errors is vital as more complex treatments, and more of them, are developed and implemented.

The work in chapter 4 that concerned gravity motion of the MLC leaves focussed on treatment plans where all of the collimator angles are 0°, i.e. where the leaves run in the up-down direction at gantry angle 90. Thus the work looks at the worst-case scenario. Although it is a worst case, it is not an unlikely case, as many of the treatment plans so far produced at Addenbrooke's have used this collimator angle. For example, head-and-neck treatments with nodal involvement require a large superior-inferior and left-right dimension to the treatment field to cover the lower head, neck and shoulder areas. To encompass such as large area with MLCs, the collimator angle needs to be at or near a cardinal angle of 0, 90, 180 or 270 degrees. Treatment planning systems typically start with a collimator angle of 0 or 180 degrees, depending on the configuration of the planning system and the linac, which may not be able to rotate the collimator to 180 degrees. If a fluence map with $1 \times 1$ cm$^2$ pixels is used, then the map will look the same at any cardinal collimator angle and there is no incentive to the treatment planner not to use the default setting. Thus a simple way to remove the

effects of gravity motion on the leaves is to use collimator angles of 90 or 270 degrees, where the MLC leaves will always run parallel to the floor and will therefore not experience a component of gravity along their direction of travel. However, the backlash in the gearing mechanisms that causes the gravity motion would manifest itself as movement hysteresis when the collimators are at 90°, leading to different leaf positions depending on the direction of travel of the leaves. This direction-dependent error could appear to be a random leaf positioning error, although it would be systematic and repeatable. In general, the MLC leaves would be too large when driven into the field and too small when driven out. If both leaf banks are driven to one side of the field, then one would be too small, one too large, and the effect would be to shift the open field to one side slightly. If, for an entire beam made up of segments, all fields are shifted in one direction, then the overall effect is to move the entire fluence map in that direction, and the matchlines and steps in fluence would not be affected. This would seem to mean that there are no rods of overdose or underdose as seen in the simulations in chapter four or when the fluence map is measured using a sheet of film. While this situation may sound advantageous to the treatment planner, perhaps it is not. With conformal radiotherapy, shifting a treatment field to the left or right is not much of a problem, as only the field edges are affected since the beam profile is flat across its width. If a fluence map is shifted across, then all the steps in fluence move across with it. While the dosimetry of a single field will not seem to be affected, the crossing points of all the fluences inside the patient will be moved, and the shifted steps in fluence will combine in different ways, leading to hot spots and cold spots as seen with the simulations shown in chapter 4.

With leaf hysteresis at non-zero collimator angles, the segmentation algorithm will make a difference to the changes in fluence maps. With a sliding window technique, where the leaves always travel from one side of the field to the other, the

situation will be as above. For a slice method, where leaves will typically move in to treat small segments (as with the pyramid shown in chapter 1), then there will be areas of over and under dosing within the fluence map itself. Thus using a collimator angle of 0 or 180 (if allowed) could be better, as although it introduces the possibility of gravity-induced motion on the leaves, this motion is at least consistent. With suitable calibration the effects of gravity motion can be offset, whereas the effects of motion hysteresis cannot easily be compensated for.

Chapter 4 has shown that a relatively simple dose calculation engine can be used to highlight potential problems with dose delivery dependent upon the exact configuration of a plan. The work has also shown that the methodology of actual dose verification needs to be carefully considered. Because the possibility of gravity motion exists for machines without direct MLC feedback such as Siemens linacs, pre-treatment dose verification needs to be carried out simulating the exact treatment conditions. In other words, it is important to use the gantry angles specified in the plan. Treatment planning systems often have a QA tool, in which they will apply the patient beam configuration to a different 'patient', typically a block of water-equivalent material. The TPS then calculates the dose distribution for this block, and measurements can be taken to compare. Treatment planning systems such as Xio and Corvus have the option when creating QA plans to reset all gantry angles to 0 (or 'nominal'). This procedure can make the actual verification easier, as planar films can be used to record the fluence maps of each beam, and the combination of all the treatment fields can be easier to check. However, this would clearly not show the effects of gravity motion or gantry-dependent MLC positioning on the dose distribution. For a Siemens-style machine with double-focussed leaves, taking all measurements at gantry angle 0 would most likely lead to a measurement that was similar to the treatment planning system dose but an underestimate of the dose the patient will receive from the real treatment. The use of

nominal gantry for QA can also be used for checking fluence maps of all the beams, as films can be placed at right angles to the beam more easily. Checking of fluence maps is often carried out to check for matchline errors as an indicator of dose delivery issues, but again at gantry 0 these will be reduced compared to measurements at other gantry angles. In order to perform thorough and accurate fluence map checks, the treatment gantry angles need to be used. Films can be used for this purpose, if mounted in the block-carrying tray of a linac's head. Alternatively, portal imagers could be used, as these rotate with the gantry and can image the radiation fields at any gantry angle.

The work carried out for chapter 4 involved the creation of several pieces of new software that were required for the research carried out. These applications have also been transferred into useful tools for the IMRT procedures currently being set up at Addenbrooke's. As with dose evaluation, any tool that assists with and speeds up IMRT work can help to encourage its use and make it more accessible to treatment planners and physicists. For example, the independent dose calculation engine, ICE, has been modified into a program called IMPoD (Intensity Modulated point Dose), which calculates the dose delivered to the isocentre and any other specified points within the patient. This program is used after the treatment planning process to check that the numbers of monitors units calculated and exported by the treatment planning are correct by comparing the dose it calculates to that originally prescribed. An independent monitor unit check is required by the guidelines laid out in the Manual of Cancer Standards (2003), and IMPoD satisfies this requirement. Other software has proved useful, such as the RTPedia software, which allows modifications to treatment files so that they can be checked pre-treatment. By modifying the file details, the treatments can be imported into the V&R system in such a way that they do not interfere with the patient's real treatment. With conformal radiotherapy treatments, with only a few fields, manual adjustments can be made, but for IMRT treatments with

around 100 segments, modifications of this nature are not feasible. An additional program has been written to assist in the importing and display of these segments within the V&R system. The RTPConnect file format allows for two different specifications of radiation fields. There is the normal definition of each field as a separate entity, which is generally used or conformal therapy, and there is a control-point definition. Control points allow for several segments of an IMRT treatment to be defined as different fixed points of one beam. For example, the first control point defines the MLCs closed with no dose, then the second point defines the leaves open, still with no dose. The third control point defines the leaves in the same position, but with the dose level raised to however many monitor units that segment needed. In this way, the V&R system and the linac move through the control points and alter aspects of the machine to reach the next point. In the simple example above, in getting from the first to the second point, the leaves need to be moved and in getting to the third point, the beam needs to be switched on until the correct number of monitor units is achieved. In this way a whole set of segments that define one fluence-modulated beam can be defined as a set of control points. Crucially, the Lantis V&R system treats such linked segments as a single multi-point entity in its interface. Thus the display in Lantis shows only 5 or 7 beam entries (i.e. the number of gantry angles) rather than 100 individual segments. Since each field needs to be manually approved before treatment can commence, this greatly reduces the workload for data input staff. Additionally, it means that individual segments cannot be accidentally left out of a treatment, and dealing with 5 entities rather than 100 clearly reduces the scope for error. Xio and Corvus export RTPConnect files as the former type, with all segments defined as individual entries. The software tool that has been created converts an RTPConnect file from this format into the control point format before import into the V&R system.

Software tools for assisting with dose verification are also vital for making the IMRT process smooth and easy to carry out. Many commercial companies sell software to assist with comparing dose measurements to calculations, many specialising in comparing sheets of film to dose planes exported from treatment planning systems. In order to assess such dose distributions using the Kappa index, new software has been created. The software has been written to make the comparison as quick and easy as possible to reduce workload on those performing the analysis. It reads a dose plane as exported by Xio, and reads isocentre co-ordinates and dose information from it. A digitized film (using a scanner such as a Vidar film digitizer) is then read by the program and is automatically aligned with the calculated dose distribution using a normalised cross-correlation. This helps to remove any possible offset-errors that might be introduced if the user is asked to manually locate the isocentre on the film. Complex dose distributions created with IMRT cannot be simply shifted in one direction by delivery errors, so using automatic alignment such as this does not remove any real errors within the dose distribution. Conversion of the digitized film to dose values (from the scanned optical densities) can use a default lookup table previously setup, or can use a cubic curve fitted to a set of calibration points taken from percentage depth-dose films or open fields delivered to films. The Kappa analysis then takes place automatically, presenting a screen with plots showing the dose difference, the distance to agreement and Kappa index plots. Also presented is a Kappa-area histogram showing the percentages of the analysis area that passes tolerance. Use of the software takes approximately 1 minute, although if all of the default options – film-dose calibration, tolerance levels and filenames – the process can be completed in around 30 seconds. It is important that tools such as this are easy and intuitive to use, as the most important part of dose verification is the analysis of the data.

The software tools that have been developed for research carried out for this thesis, and the spin-off tools created alongside, have aided and helped to speed up the implementation of IMRT at Addenbrooke's. During the course of this thesis, the tools have helped in the validation of IMRT and the delivery processes, which would not have been possible without software assistance. Thus the work presented here and the knowledge gained from it have easily transferred into clinical relevance. The results of this research have helped to guide the limits and tolerances that are being implemented for IMRT at Addenbrooke's.

Many of the concepts in this thesis focus on the absolute physical aspects of radiotherapy, such as patient position, dose difference and distance to agreement. However, the aim of the research is to improve not only the accuracy (which very much has a physical meaning) but the *outcome* of radiotherapy. Chapters 5 and 6 dealt with the radiobiological effects of radiotherapy. The work presented in chapter 5 shows that there is no implicit equivalence between the numerical values of millimetres distance to agreement and percentage dose difference when analysing dose distributions.

In order to test the equivalence or otherwise of distance to agreement and dose difference, the equivalent uniform dose was used. This measure is a useful tool as it enable the conversion of a shifted dose distribution to an equivalent dose and therefore an equivalent dose difference, which then allows comparison with an actual dose difference. The work shows that for typical target volumes, which tend to have high uniform doses across them, a physical movement does not equate to as high a dose difference as the penumbra and dose gradients would suggest. Conversely for organs at risk, certain movements, especially those towards the target volume, result in much higher than expected dose changes. These results question how physical measures of dose agreement are used. The respective levels of distance and dose tolerance must be

carefully chosen if physical dose evaluation is used. Alternatively, the use of EUD allows us to take a step away from the actual dose distribution and look at the equivalent delivery to the patient. Certainly distance to agreement is less useful a measure for IMRT than for conventional radiotherapy. With conventional radiotherapy, any distance differences will manifest themselves at the edges of the treatment area, where treatment planning margins are added anyway to ensure target coverage. But with IMRT, doses across the target volume are not delivered using uniform fields and thus there are likely to be errors within the target volume where there have been and can be no margins applied.

Chapter 6, then, moves on to determine how changes in the accuracy of a treatment affect the outcome of the treatment for the patient. Two radiobiological tools are used, the EUD and the NTCP, leading to the probability of uncomplicated control and the therapeutic gain (as discussed in chapter 1). The results show that when the EUD is used as an analysis tool, the dosimetric changes in a patient treatment are larger than those predicted by composite indices such as Gamma and Kappa, despite Kappa being a relatively intolerant measure when compared to Gamma. The work also shows that composite indices, although presenting a numerical index, do not tell us much about the nature of any dose differences encountered in terms of what that difference actually means. A more effective measure of dose distribution accuracy could therefore be the EUD as it shows the overall effect of an inhomogeneous dose distribution in terms of dose, rather than a slightly confusing amalgam of distance and dose. After all, an oncologist prescribes a dose to a tumour and organs at risk, and is aware of the acceptable differences he or she will accept. The EUD can present the effective dose to a structure as a single number which is informative. If a dose distribution passes a dose/distance composite index with a dose tolerance of 3%, it may not equate to the same overall change in EUD. Composite indices separate the dose distribution from the

reality that a patient is going to be treated with it whereas EUD brings into the frame some radiobiological modelling and refocuses the analysis on the patient themselves.

Building on this, chapter 6 developed the EUD measurements into estimates of patient outcomes in terms of the probabilities of tumour control and normal tissue complication. The work shows that the probability of uncomplicated tumour control and the therapeutic gain both change as the accuracy of the dose delivery changes in ways explained in chapter 4. Surprisingly though, the results show that the best patient outcome is *not* associated with the treatment delivered as planned, but when there are inaccuracies in the treatment machine leading to larger treatment fields. These results merely highlight the fact that physical measures of dose difference are far removed from the reality of dose delivery to a living patient. The results of the PUC and TG calculations clearly show that the best outcome for a patient is achieved when the composite distance/dose indices would indicate a failure of the dose distribution to meet tolerances set upon it.

Many of the concepts and treatment of errors in this thesis have important implications for the QA framework a radiotherapy centre employs to ensure that patients are treated accurately and well. Many of these implications centre around pre-treatment verification of IMRT dose distributions. Chapter 3 shows that the different dose evaluation metrics available, although perhaps claiming to be the same as the Gamma index, treat dose gradients and rapidly changing dose levels in different ways, leading to some metrics being much more tolerant than others. Clearly a firm understanding of how an evaluation metric works is crucial when choosing to use it for the purposes of IMRT verification.

Chapter 4 demonstrates that many of the errors that can affect IMRT dose distribution in a serious way are systematic and with a little investigation of the

properties of a treatment machine most can be uncovered and compensated for. Crucially, the work shows that any dose verification work *must* use the same machine settings as those used for treatment. Linear reduction of monitor units to make verification runs faster, for example, will alter the dose linearity with which the machine delivers the radiation. Setting all beam angles to gantry 0, also to make the verification easier and faster, is also going to alter the way the dose is delivered as any errors, systematic or random, associated with gantry rotation will not be including in the verification. Chapter 4 shows that even small systematic errors in MLC leaf position can lead to potentially large errors in dose delivery, and that these errors would not necessarily be picked up if the patient treatment is not delivered during verification.

Chapters 5 and 6 question how the verification itself is carried out. Much of what treatment planners and physicists do is unrelated to the patient directly, in that doses and tolerances are considered without the needs and outcome of the patient necessarily incorporated. Chapter 5 shows that the common tolerances used in physical dose verification may need to be rethought because millimetres of movement do not correspond to percentage dose differences. Additional to this, chapter 6 shows that when the results of verification are taken right through to the endpoint of patient delivery, the probabilistic outcome of the treatment is not linked to the accuracy of the delivery in an intuitive way. In terms of treatment verification, this implies that every verification needs to work results all the way through radiobiological models in order to asses whether or not the verification that is carried out is suitable, and also to consider if the results of a 'bad' verification are necessarily bad for the patient. Conversely to the verification process, the work here has further implications for the treatment planning process. Dose differences such as those simulated in chapter 4 actually appear to be better for the patient. It has to be assumed that the dose distribution as planned was believed by the prescribing oncologist and the treatment planner to be the best for the

patient. The fact that a radiobiological analysis has implied that a different dose distribution, arrived at by random chance of a systematic error, is better for the patient shows that there may be room for improvement in treatment planning. Incorporation of radiobiological modelling into treatment planning is currently available in many treatment planning systems, but is not widely used. The results of this thesis show that there may be significant gains involved in radiotherapy in implementing modelling into the treatment planning process. Surprising and non-intuitive results may result from a treatment planning process guided by patient outcome, but if the statistics and variables are accurately determined beforehand, then there is every reason to believe that the results will translate into better patient care.

Although the work in chapter 6 implies that patient outcome is more important that accuracy of treatment delivery, it is still vitally important to be aware of errors in the treatment process so that they can either be corrected for or properly modelled. Even though certain errors in treatment delivery may lead to improvements for the patient, we cannot know this if we do not know where or what those errors are. Incorporation and modelling of such errors requires a fully electronic solution, such as the ICE dose calculation engine or a modifiable treatment planning system. It is not possible to measure dose distributions inside patients, although there are systems that allow exit fluence to be analysed to calculate the dose deposited. Measurements of dose distributions inside phantom materials can be informative in highlighting potential errors in the delivery process, but cannot translate them to the patient, as the phantom is a different material and shape from the patient and will have a different dose distribution affected in different ways. Therefore in-house software such as ICE can be useful in predicting the likely changes in dose, and can be easily changed to incorporate radiobiological models directly into the analysis of dose differences. The ultimate goal

would be to have treatment planning systems that can incorporate known treatment errors such that they can be included in the optimisation of the treatment plan. This would eliminate the need for external software such as ICE and also add an extra degree of confidence that what the treatment planning system calculates is what will be delivered to the patient. The planning system will then be able to produce estimates of patient outcome based on the actual dose distribution.

Even if modifiable treatment planning systems and full radiobiological modelling is not possible, knowledge of errors is still important, as they can be incorporated into prescription of dose at an early stage in the treatment planning process. For example, if analysis shows that a certain type of treatment plan is likely to give dose errors equivalent to 5 Gray EUD to the spinal cord, then a treatment planning tolerance level for that organ can be set 5 Gray below what the oncologist believes is acceptable, therefore ensuring that the organ does not receive a dose higher than what is acceptable. In this way, dose and dose-error analysis can be fed back into treatment planning processes, which while not necessarily improving the accuracy of treatments, can go some way towards ensuring that outcomes are improved or at the very least not jeopardised.

# Chapter 8 – Conclusions

This thesis has evaluated and questioned several key concepts in assessing the accuracy of IMRT treatments.

In order to make improvements in accuracy, a radiotherapy centre needs to know the current level. A non-invasive method for assessing patient setup accuracy without the use of radiation has been described, which has identified the best immobilisation system to use for head and neck patients at Addenbrooke's radiotherapy centre. The technique involves the use of a scanning laser system, which is relatively mobile. Therefore centres which do not have access to such a system may be able to enter into regional collaboration at little expense in order to assess their own levels of accuracy in patient setup.

In terms of dose distributions, accuracy must be assessed using evaluation tools, and several composite methods which combine distance and dose measured were evaluated. It was found that although several published methods claim to implement the same concept, their actual tolerance levels differ. A new method is proposed for accurately calculating the Gamma index in the way that it was originally intended. Additionally, a less tolerant evaluation metric is proposed, such that potential problem areas in dose distributions can be highlighted where they might pass the more tolerant evaluations.

The new dose evaluation metric was used to analyse dose differences created by systematic treatment machine errors, and the results showed that large delivery errors could result from small changes in machine behaviour. Positional and dosimetric errors were simulated using an in-house dose calculation engine which showed that treatment machines must be well calibrated in order to achieve good treatment accuracy.

Many implementations of composite dose evaluation indices use the same value for percentage dose difference and millimetres distance tolerances. The relative levels of dose and distance tolerance were evaluated and found not to be numerically equivalent. Suggestions have been made for more suitable relative levels when composite analyses are used.

Building on the concept of improving dose evaluation, the dose difference found from simulation of machine characteristics were evaluated in terms of patient outcome. The results show that the patient outcome is affected by the dose differences but in non-intuitive ways.

It is suggested that dose verification methods be reviewed in terms of what is analysed and how the results are treated. This work shows that there may be assumptions about IMRT treatments that may jeopardise patient care in unpredictable and complex ways. Treatment planning systems should implement radiobiological modelling as one of the primary parameters on which treatment plans are optimised and calculated and they should also be modifiable in order that sources of potential error can be incorporated into them. These changes would allow accurate determination of the level of dose that a patient will receive during treatment, and also show the effects of that dose.

IMRT is a complex procedure and in order to ensure that patient care is as good as we can make it, the entire treatment chain needs to be scrutinised to determine where errors lie and the likely effects these errors have. This is particularly crucial for IMRT as the goals of conventional radiotherapy are being arrived at via a different route, and many of the processes and checking that are employed have not changed from those of conventional radiotherapy. IMRT is being implemented as an improvement to patient care, so it should be ensured that the processes involved are indeed improving patient care. Many errors specific to IMRT have been highlighted in this thesis, but this work

is not a complete toolkit for the improvement of IMRT. The work presented here shows several ways in which the accuracy of IMRT can be measured and therefore improved. Building on this work, and running parallel to it, methods for analysing patient outcome have been presented, which then allows for the improvement of patient care to be implemented.

# References


Abramowitz M and Stegun IA (1968) *Handbook of mathematical functions*, Washington, DC: National Bureau of Standards

Agazaryan N, Solberg T and DeMarco J (2003) "Patient specific quality assurance for the delivery of intensity modulated radiotherapy" *J. Appl. Clin. Med. Phys.* **4**, 40–50

Ambrose J and Hounsfield GN (1973), "Computerized transverse axial tomography", *Brit J Radiol* **46** (542), 148-149

Bakai A, Alber M and Nüsslin F (2003), "A revision of the γ-evaluation concept for the comparison of dose distributions", *Phys Med Biol* **48**(21), 3543-3553

Beavis AW, Weston SJ and Whitten VJ (1996), "Implementation of the Varian EDW into a commercial RTP system", *Phys Med Biol*, **41**(9), 1691-1704

Beavis AW (2001), "Optimization of the step-and-shoot leaf sequence for delivery of intensity modulated radiation therapy using a variable division scheme", *Phys Med Biol* **46**(9), 2457-2465

Beaver M, Matheny K, Roberts D and Myers J (2001), "Predictors of weight loss during radiation therapy", *Otolaryngology-Head And Neck Surgery* **125**(6), 645-648

BIR (2003), *Geometric Uncertainties in Radiotherapy. Defining the Planning Target Volume*, British Institute of Radiology Publishing

Bley CR, Blattman H, Roos M, Sumova A and Kaser-Hotz B (2003), "Assessment of a radiotherapy patient immobilization device using single plane port radiographs and a remote computed tomography scanner", *Veterinary Radiology & Ultrasound* **44**(4), 470-475

Bortfeld T, Kahler D, Waldron T and Boyer A (1994), "X-ray field compensation with multileaf collimators", *Int J Rad Onc Biol Phys* **28**(3), 723-730

Boyer AL (1992), Antonuk L, Fenster A van Herk M, Meetens H, Munro P, Reinstein L and Wong J, " A review of electronic portal imaging devices (EPIDS)", *Med Phys* **19**(1), 1-16

Budgell GJ, Mot JHL, Williams PC and Brown KJ (2000), "Requirements for leaf position accuracy for dynamic multileaf collimation", *Phys Med Biol* **45**(5), 1211-1227

Burman C, Kutcher GJ, Emami B and Goitein M (1991), "Fitting of normal tissue tolerance data to an analytic function", *Int J Rad Onc Biol Phys* **21**(1), 123-135

Carlson H (2001), "Intensity modulation using multileaf collimators: Current status", *Med Dosim* **26**, 151-156

Cernica G, de Boer SF, Diaz A, Fenstermaker RA and Podgorsak MB (2005), "Dosimetric accuracy of a staged radiosurgery treatment", *Phys Med Biol* **50**(9), 1991-2002

Chapet O, Thomas E, Kessler M, Fraass B and Ten Haken R (2005), "Esophagus sparing with IMRT in lung tumour irradiation: An EUD-based optimization technique", *Int J Rad Onc Biol Phys* **63**(1), 179-187

Childress NL and Rosen II (2003), "The design and testing of novel clinical parameters for dose comparison", *Int J Rad Onc Biol Phys* **56**(5), 1464-1479

Cheng C-W and Das IJ (1996), "Dosimetry of high energy photon and electron beams", *Med Phys* **23**(7), 1225-1232

Chui CS, Spirou SV and LoSasso T (1996), "Testing of dynamic multileaf collimation", *Med Phys* **23**(5) 635-641

Chung HT, Jin HS, Dempsey JF, Liu CH, Palta J, Suh TS and Kim SY (2005), "Evaluation of surface and build-up region dose for intensity-modulated radiation therapy in head and neck cancer", *Med Phys* **32**(8), 2682-2689

Convery D and Rosenbloom M (1992), "The generation of intensity-modulated fields for conformal radiotherapy by dynamic collimation", *Phys Med Biol* **37**(6), 1359-1374

Dahlin H, Lamm IL, Landberg T, Levernes S and Ulso N (1983), "User requirements on CT based computerized dose planning systems in radiotherapy", *Acta Radiol: Oncol* **22**, 397-415

Dearnaley DP, Khoo VS, Norman AR, Meyer L, Nahum A, Tait D, Yarnold J and Horwich A (1999), "Comparison of radiation side-effects of conformal and conventional radiotherapy in prostate cancer: a randomised trial", *The Lancet* **353**(9149), 267-272

De Boer HCJ, van Sörnsen de Koste JR, Creutzberg CL, Visser AG, Levendag PC and Heijman BJM (2001), "Electronic portal image assisted reduction of systematic set-up errors in head and neck irradiation", *Rad & Onc* **61**, 299-308

Depuydt T, Van Esch A, Huyskens DP (2002), "A quantitative evaluation of IMRT dose distributions: refinement and clinical assessment of the Gamma evaluation", *Rad & Onc* **62**, 309-319

Duthoy W, De Gersem W, Vergote K, Boterberg T, Derie C, Smeets P, De Wagter C and De Neve W (2004), "Clinical implementation of intensity-modulated arc therapy (IMAT) for rectal cancer", *Int J Rad Onc Biol Phys* **60**(3), 794-806

Ebert M (2000), "Viability of the EUD and TCP concepts as reliable dose indicators", *Phys Med Biol* **45**(2), 441-457

Emami B, Lyman J, Brown A, Coia L, Goitein M, Munzenrider JE, Shank B, Solin LJ and Wesson M (1991), "Tolerance of normal tissue to therapeutic irradiation", *Int J Rad Onc Biol Phys* **21**(1), 109-122

Freund L (1897), "Ein mit Röntgen-Strahlen behandelter Fall von Naevus pigmentosus piliferus", *Wien Med Wschr* **47**, 428-434

Fuss M, Salter BJ, Cheek D, Sadeghi A, Hevezi JM abd Herman TS (2004), "Repositioning accuracy of a commercially available thermoplastic mask system", *Rad & Onc* **71**, 339-345

Gilbeau L, Octave-Prignot M, Loncol T, Renard L, Scalliet P and Grégoire V (2001), "Comparison of setup accuracy of three different thermoplastic masks for the treatment of brain and head and neck tumours", *Rad & Onc* **58**, 155-162

Gilhuijs K, van de Ven P and van Herk M (1996), "Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data", *Med Phys* **23**(3), 389-399

Greer PB, Mortensen TM, Rad DT and Jose CC (1998), "Comparison of two methods for anterior-posterior isocenter localization in pelvic radiotherapy using electronic portal imaging", *Int J Rad Onc Biol Phys* **41**(5), 1193-1199

Hanna CL, Slade S, Mason MD and Burnet NG (1999), "Accuracy of patient positioning during radiotherapy for bladder and brain tumours", *Clinical Oncology* **11**(2), 93-98

Harms WB, Low DA, Wong JW and Purdy JA (1998), "A software tool foe the quantitative evaluation of 3D dose calculation algorithms", *Med Phys* **25**(10), 1830-1836

Haus AG, Dickerson RE, Huff KE, Monte S, Schlager BA, Atanas M and Matloubieh A (1997), "Evaluation of a cassette-screen combination for radiation therapy portal localization imaging with improved contrast", *Med Phys* **24**(10), 1605-1608

Hawkes DJ (1998), "Algorithms for radiological image registration and their clinical application", *J Anat* **193**, 347-361

Hess CF, Kortmann RD, Jany R, Hamberger A and Bamberg M (1995), "Accuracy of field alignment in radiotherapy of head and neck-cancer utilizing individualized face mask immobilization - a retrospective analysis of clinical-practice", *Rad & Onc* **34**(1), 69-72

Hogstrom KR, Mills MD, Meyer JA, Palta JR, Mellenberg DE, Meoz RT and Fields RS (1984), "Dosimetric evaluation of a pencil-beam algorithm for electrons employing a two-dimensional heterogeneity correction", *Int J Rad Onc Biol Phys* **10**(4), 561-569

Holthusen H (1936), "Erfahrungen über die Verträglichkeitsgrenze für Röntgenstrahlen und deren Nutzanwendung zur Verhütung von Schäden", *Strahlentherapie* **57**, 254-269

Hounsell A and Jordan T (1997), "Quality control aspects of the Philips multileaf collimator", *Rad & Onc* **45**, 225-233

Hounsfield GN (1973), "Computerized transverse axial scanning (tomography). 1. Description of system", *Brit J Radiol* **46** (552), 1016-1022

Humphreys M, Urbano MTG, Mubata C, Mites E, Harrington KJ, Bidmead M and Nutting CM (2005), "Assessment of a customised immobilisation system for head and neck IMRT using electronic portal imaging", *Rad & Onc* **77**(1), 39-44

Huq MS, Das IJ, Steinberg T and Galvin JM (2002), "A dosimetric comparison of various multileaf collimators", *Phys Med Biol* **47**(12), N159-N170

ICRU (1987), "Report 42: Use of computers in external beam radiotherapy procedures with high-energy photons and electrons"

ICRU (1993), "Report 50: Prescribing recording and reporting photon beam therapy"

ICRU (1999) – Report 62: Supplement to report 50

Jobse M, Davelaar J, Hendriks E, Kattevilder R, Reiber H and Stoel B (2003), "A new algorithm for the registration of portal images to planning images in the verification of radiotherapy, as validated in prostate treatments", *Med Phys* **30**(9), 2274-2281

Jordan T and Williams PC (1994), "The design and performance characteristics of a multileaf collimator", *Phys Med Biol* **39**(2), 231-251

Kijewski PK, Chin LM and Bjarngard BE (1978), "Wedge-shaped dose distributions by computer-controlled collimator motion", *Med Phys* **5**(5), 426-429

Langmack K (2001), "Portal Imaging", *Brit J Radiol* **74**, 789-804

Leavitt DD, Martin M, Moeller JH and Lee WL (1990), "Dynamic wedge field techniques through computer-controlled collimator motion and dose delivery", *Med Phys* **17**(1), 87-91

Levene MB, Kijewski PK, Chin LM, Bjarngard BE and Hellman S (1978), "Computer-controlled radiation-therapy", *Radiology* **129**(3), 769-775

Li XA, Wang JZ, Stewart RD and DiBiase SJ (2003), "Dose escalation in permanent brachytherapy for prostate cancer: dosimetric and biological considerations", *Phys Med Biol* **48**(17), 2753-2765

LoSasso T, Chui CS and Ling CC (1998), "Physical and dosimetric aspects of a multileaf collimation system used in dynamic mode for implementing intensity modulated radiotherapy", *Med Phys* **25**(10), 1919-1927

Low DA and Dempsey JF (2003), "Evaluation of the Gamma dose distribution comparison method", *Med Phys* **30**(9), 2455-2464

Low DA, Harms WB, Mutic S and Purdy JA (1998), "A technique for the quantitative evaluation of dose distributions", *Med Phys* **25**(5), 656-661

Mackie T, Holmes T and Swerdloff S (1993), "Tomotherapy – A new concept for the delivery of dynamic conformal radiotherapy", *Med Phys* **20**(6), 1709-1719

*Manual of Cancer Services Standards* (2003), NHS Executive

McKenzie AL, van Herk M and Mijnheer B (2000), "The width of margins in radiotherapy treatment plans", *Phys Med Biol* **45**(11), 3331-3342

Meedt G, Alber M and Nusslin F (2003), "Non-coplanar beam direction optimization for intensity-modulated radiotherapy", *Phys Med Biol* **48**(18), 2999-3019

Miften M et al (2000), "Implementation and verification of virtual wedge in a three-dimensional radiotherapy planning system", *Med Phys* **27**(7), 1635-43

Munro P (1995), "Portal imaging technology: Past present and future", *Semin Radiat Oncol* **5**, 115-133

Niemierko A (1997), "Reporting and analyzing dose distributions: A concept of equivalent uniform dose", *Med Phys* **24**(1), 103-110

Niemierko A (1999), "A generalized concept of equivalent uniform dose (EUD)", *Med Phys* **26**(6), 1100

Olch AJ (2002), "Dosimetric performance of an enhanced dose range radiographic film for intensity-modulated radiation therapy quality assurance", *Med Phys* **29**(9), 2159-2168

Paliwal B, Tomé W, Richardson S and Mackie TR (2000), "A spiral phantom for IMRT and tomotherapy treatment delivery verification", *Med Phys* **27**(11), 2503-2507

Palta J and Mackie T (Eds) (2003), *Intensity modulated radiation therapy: The state of the art*, Medical Physics Publishing

Plane JH (1985), "Comparison of radiographic techniques of treatment alignment in radiotherapy", *Radiography* **51**(598), 211-214

Prisciandaro JI, Frechette CM, Herman MG, Brown PD, Garces YI and Foote RL (2004), "A methodology to determine margins by EPID measurements of patient setup variation and motion as applied to immobilization devices", *Med Phys* **31**(11), 2978-2988

Que W, Kung J and Dai J (2004), "'Tongue-and-groove' effect in intensity modulated radiotherapy with static multileaf collimator fields", *Phys Med Biol* **49**(2), 399-405

Ramsey C, Spencer K, Alhakeem R and Oliver A (2001), "Leaf position error during conformal dynamic arc and intensity modulated arc treatments", *Med Phys* **28**(1), 67-72

Richardson SL, Tomé WA, Orton NP, McNutt TR and Paliwal BR (2003), "IMRT delivery verification using a spiral phantom", *Med Phys* **30**(9), 2553-2558

Sharp L, Lewin F, Johansson H, Payne D, Gerhardsson A and Rutqvist LE (2005), "Randomized trial on two types of thermoplastic masks for patient immobilization during radiation therapy for head-and-neck cancer", *Int J Rad Onc Biol Phys* **61**(1), 250-256

Siochi RAC (1999), "Minimizing static intensity modulation delivery time using an intensity solid paradigm", *Int J Rad Onc Biol Phys* **43**(3), 671-680

Stasi M, Giordanengo S, Cirio R, Boriano A, Bourhaleb F, Cornelius I, Donetti M, Garelli E, Gomola I, Marchetto F, Porzio M, Sanz Freire CJ, Sardo A and Peroni C (2005), "D-IMRT verification with a 2D pixel ionization chamber: dosimetric and clinical results in head and neck cancer", *Phys Med Biol* **50**(19), 4681-4694

Stell A, Li J, Zeidan O and Dempsey J (2004), "An extensive log-file analysis of step-and-shoot intensity modulated radiation therapy segment delivery errors", *Med Phys* **31**(6), 1593-1602

Stock M, Kroupa B and Dietmar G (2005), "Interpretation and evaluation of the γ index and the γ angle for the verification of IMRT hybrid plans", *Phys Med Biol* **50**(3), 399-411

Suit H (2002), "The Gray Lecture 2001: Coming technical advances in radiation oncology", *Int J Rad Onc Biol Phys* **53**(4), 798-809

Sweeney R, Bale R, Vogele M, Nevinny-Stickel M, Bluhm A, Auer T, Hessenberger G and Lukas P (1998), "Repositioning accuracy: Comparison of a non-invasive head older with thermoplastic mask for fractionated radiotherapy and a case report", *Int J Rad Onc Biol Phys* **41**(2), 475-483

Terahara A, Niemierko A, Goitein M, Finkelstein D, Hug E, Liebsch N, O'Farrell D, Lyons S and Munzenrider J (1999), "Analysis of the relationship between tumor dose inhomogeneity and local control in patients with skull base chordoma", *Int J Rad Onc Biol Phys* **45**(2), 351-358

Thomas S J and Brooke S (2002), "Isocentre blurring: a solution to the limitations of the PTV concept in IMRT", *Int J Rad Onc Biol Phys* **54**(2), 324

Tsai J-S, Engler MJ, Ling MN, Wu JK, Kramer B, Dipetrillo T and Wazer DE (1999), "A non-invasive immobilization system and related quality assurance for dynamic intensity modulated radiation therapy of intracranial and head and neck disease", *Int J Rad Onc Biol Phys* **43**(2), 455-467

Van Dyk J, Barnett RB, Cygler JE and Shragge PC (1993), "Commissioning and quality assurance of treatment planning computers", *Int J Rad Onc Biol Phys* **26**(2), 261-273

Van Herk M, Remeijer P, Rasch C and Lebesque JV (2000), "The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy", *Int J Rad Onc Biol Phys* **47**(4), 1121-35

Van Santvoort J (1998), "Dosimetric evaluation of the Siemens Virtual Wedge.", *Phys Med Biol* **43**(9), 2651-63

Waite K and Filshie J (1990), "The use of a laryngeal mask airway for CT radiotherapy planning and daily radiotherapy", *Anaesthesia* **45**(10), 894-894

Webb S (2001), *Intensity-Modulated Radiation Therapy*, IOP Publishing

Webb S and Lomax T (2001), "There is no IMRT?", *Phys Med Biol* **46**(12), L7-L8

Weltens C, Kesteloot K, Vandevelde G and Van den Bogaert W (1995), "Comparison of plastic and Orfit® masks for patient head fixation during radiotherapy: Precision and costs", *Int J Rad Onc Biol Phys* **33**(2), 499-507

Willner J, Hädinger U, Neumann M, Schwab FJ, Bratengeier K and Flentje M (1997), "Three dimensional variability in patient positioning using bite block immobilization in 3D-conformal radiation treatment for ENT-tumors", *Rad & Onc* **43**, 315-321

Woo K and Nico A (2005), "Impact of multileaf collimator leaf positioning accuracy on intensity modulation radiation therapy quality assurance ion chamber measurements", *Med Phys* **32**(5), 1440-1445

Zeidan O, Li J, Ranade M, Stell A and Dempsey J (2004), "Verification of step-and-shoot IMRT delivery using a fast video-based electronic portal imaging device", *Med Phys* **31**(3), 463-476

APPENDIX NOT COPIED

ON INSTRUCTION FROM

UNIVERSITY